Characterization and Development of a Novel Cytomegalovirus for Evaluation as an HIV Vaccine Viral Vector in a Non-Human Primate Model

by

Angie K. Marsh

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Immunology
University of Toronto

© Copyright by Angie Marsh 2014
Characterization and Development of a Novel Cytomegalovirus for Evaluation as an HIV Vaccine Viral Vector in a Non-Human Primate Model

Angie Marsh
Doctor of Philosophy
Department of Immunology
University of Toronto
2014

Abstract

Within the field of HIV vaccine design, it has been recognized that there is a real need to diversify the pool of vectors undergoing testing and to utilize vectors that can induce durable immunity. Herpesviruses, such as cytomegalovirus (CMV), show much promise as viral vectors, as they establish lifelong latency with periodic reactivation, enabling persistent expression of vaccine antigens. We propose that using CMV as a lifelong reacting HIV viral vector will generate a robust and persistent effector memory T cell response at the mucosa and thus, may impair HIV replication at its earliest stage. In an effort to evaluate CMV-based HIV/SIV vectors in a non-human primate (NHP) model, I isolated CMV from cynomolgus macaques (CyCMV), and performed preliminary phenotypic and genomic classification of the novel virus. To characterize the CyCMV genome, facilitate targeted recombination for vaccine construction, and inform possible attenuation strategies, I sequenced and annotated the complete viral genome using next-generation sequencing. Furthermore, I compared and contrasted the structural and functional genes of CyCMV to human and rhesus macaque CMV with respect to genes involved in pathogenesis, immune evasion, species-specificity, and more. In order to maximize the ability to manipulate the virus for vaccine purposes, I cloned CyCMV as a Bacterial Artificial
Chromosome (BAC). The cloning of CyCMV-BAC enabled the virus to be transformed into a bacterial system for downstream insertion of vaccine antigens, as well as future viral manipulations, including attenuations. This newly characterized CMV provides a novel model in which to study HIV/SIV CMV-based vaccines in hopes of informing forthcoming human clinical trials. Additionally, CyCMV-BAC will provide an invaluable resource for CMV biologists allowing for virus manipulation studies, functional characterization of CMV genes, as well as HCMV pathogenesis and vaccine development research.
Acknowledgments

I would like to begin by thanking my supervisor, Dr. Kelly MacDonald, for your guidance, support, and for training me in translational research with both your scientific and clinical expertise. You have always trusted me to design experiments and conduct my research independently, which has significantly fostered my growth as a PhD student. I am grateful for the incredible experiences you have given me, including multiple opportunities to attend International conferences, and for always encouraging me and trusting me to present our research at these forums. Beyond my scientific training, you have provided me with insight into career development, and managing professional relationships, which will be of great value throughout my career. To my co-supervisor, Dr. David Willer, I cannot thank you enough for your scientific mentorship and friendship, I could not have done this without you. You have passed on a wealth of scientific knowledge, and given me invaluable guidance through my studies with respect to technical training, study design, scientific writing, oral presentations, and scientific integrity. You are one of the most generous, helpful, and welcoming people I have ever met, and in addition your scientific training, you have given me great life advice throughout the years, which has included everything from helpful career advice, delicious food recipes, great drinks, new hobbies/sports, to where to plan my next vacation destination! To my virology mentor, Dr. Aruna Ambagala, I want to thank you for being an excellent teacher, role model, and friend, and most importantly for continuing to challenge me scientifically even after leaving the lab. I am indebted to you for passing on your scientific knowledge, technical skills, and most of all your strong scientific principles, there was no greater honour than knowing I gained your trust as a scientist! Jacqueline Chan, technician extraordinaire, you are one of the few people that I have worked with and had morning coffee with throughout my entire PhD. You are an excellent teacher, a skilled technician, and your organization skills are the reason my experiments were completed in a timely manner, without contamination, and with success. I cannot thank you enough for answering all of my questions, for your technical training, and for maintaining the lab to the highest standards. To my fellow MacDonald lab graduate students, Catia and Justen, it has been a pleasure having you as peers even though it was short and sweet. If it weren’t for you both I would never know the feeling of being a part of an award-winning centerpiece! Thank you to past and present members of the MacDonald laboratory, Dr. Rosemarie Mason, Dr. Oluwadamilola Iwajomo, and Dr. Kamran Haq for their support and guidance. To Olena
Skokovets, and Robbyn Sargent, I am grateful to both of you for your research contributions during your co-op placements. To the members of the Kaul and Ostrowski laboratories, thank you for taking me in as one of your own, for being a constant support network for both scientific and social discussion, and most of all for being a great group of friends.

I would like to thank my committee members, Dr. Lori Frappier and Dr. Rupert Kaul, for their ongoing support and direction. I have sincerely appreciated your constructive feedback, guidance, and your strong commitment to keeping me on track. You have assisted in keeping me focused on my project objectives in order to complete my PhD in the proposed timeframe and I thank you for that. To everyone in the Department of Immunology, thank you for being such a welcoming and social group of individuals. It has been an incredible experience, and I am honoured to have been a part of this great Department filled with bright graduate students, helpful administration, and a faculty of Internationally recognized investigators with a diverse array of expertise.

And last but certainly not least, to my parents, my brother, and Konrad, thank you for your constant support and encouragement through this endeavor, I would not be where I am today without all of you. In addition, I want to thank my extended family and friends for always being there and for cheering me on along the way.
# Table of Contents

Acknowledgments......................................................................................................................... iv

Table of Contents ............................................................................................................................ vi

List of Tables ........................................................................................................................................ xi

List of Figures ....................................................................................................................................... xii

List of Abbreviations .......................................................................................................................... xv

List of Appendices .............................................................................................................................. xx

List of Manuscripts Arising From This Thesis .................................................................................... xxi

Chapter 1 ............................................................................................................................................... 1

1 Introduction ......................................................................................................................................... 1

1.1 Human Immunodeficiency Virus ................................................................................................. 1

1.1.1 HIV/SIV Structure and Genome ............................................................................................. 3

1.1.2 HIV Replication Cycle ........................................................................................................... 6

1.1.3 HIV-1 Pathogenesis and Infection .......................................................................................... 8

1.1.4 Immune Response to HIV Infection ....................................................................................... 16

1.2 HIV/SIV Vaccine Development .................................................................................................... 22

1.2.1 Obstacles in the Development of an HIV Vaccine ................................................................. 24

1.2.2 HIV/SIV Vaccine Strategies .................................................................................................. 26

1.2.3 HIV Vaccine Clinical Trials .................................................................................................. 30

1.3 Herpesviruses as Viral Vectors for an HIV/SIV Vaccine ............................................................. 38

1.3.1 Rationale for Herpesvirus-Based Viral Vectors ..................................................................... 38

1.3.2 Herpesviruses as Viral Vectors ............................................................................................. 42

1.3.3 Cytomegalovirus .................................................................................................................... 45

1.3.4 Overarching Hypothesis ........................................................................................................ 54

1.4 Non-Human Primate Models ........................................................................................................ 55
1.5 Thesis Background and Objectives............................................................................... 60

1.5.1 Species-Specificity of Non-Human Primate Cytomegaloviruses .................. 60

1.5.2 Thesis Objectives .................................................................................................. 63

Chapter 2 ......................................................................................................................... 64

2 Isolation and Characterization of Cynomolgus Macaque Cytomegalovirus (CyCMV) .... 64

2.1 Chapter Overview ..................................................................................................... 65

2.2 Abstract ..................................................................................................................... 66

2.3 Introduction ................................................................................................................ 67

2.4 Material and Methods ................................................................................................. 68

2.4.1 CyCMV Isolation .................................................................................................. 68

2.4.2 Viral DNA Isolation ............................................................................................... 68

2.4.3 Transmission Electron Microscopy .......................................................................... 69

2.4.4 Polymerase Chain Reaction Amplification .......................................................... 70

2.4.5 CyCMV Growth Kinetics ...................................................................................... 70

2.4.6 Western Blot .......................................................................................................... 71

2.5 Results ......................................................................................................................... 71

2.5.1 Isolation of CMV from Mauritius Cynomolgus Macaques ............................. 71

2.5.2 Ultrastructural Morphology of CyCMV .............................................................. 72

2.5.3 Gross Viral Genome Analysis ............................................................................... 73

2.5.4 Phylogenetic Analysis of Conserved Genes ......................................................... 74

2.5.5 Growth Kinetics of CyCMV ................................................................................ 76

2.5.6 MHC Class I Downregulation ............................................................................. 78

2.6 Discussion .................................................................................................................... 79

Chapter 3 .......................................................................................................................... 83

3 Genomic Sequencing and Characterization of Cynomolgus Macaque Cytomegalovirus ..... 83

3.1 Chapter Overview ..................................................................................................... 84
3.2 Abstract ......................................................................................................................... 85
3.3 Introduction ................................................................................................................... 86
3.4 Materials and Methods ............................................................................................... 87
   3.4.1 CyCMV Viral DNA Isolation ................................................................................. 87
   3.4.2 Next-Generation DNA Sequencing ....................................................................... 88
   3.4.3 Bioinformatic Assembly ......................................................................................... 88
   3.4.4 Gap Closing ........................................................................................................... 89
   3.4.5 PCR Sequencing .................................................................................................... 89
   3.4.6 Error Correction .................................................................................................... 89
   3.4.7 Open Reading Frame Assignment ......................................................................... 90
   3.4.8 Accession Number ............................................................................................... 90
3.5 Results .......................................................................................................................... 90
   3.5.1 Genomic Analysis of CyCMV .............................................................................. 90
   3.5.2 Restriction Enzyme Digestion .............................................................................. 91
   3.5.3 Gene Assignment .................................................................................................. 92
   3.5.4 Genes Missing From Other Macaque CMVs ......................................................... 95
   3.5.5 Tropism Genes .................................................................................................... 96
   3.5.6 Functional genes .................................................................................................. 96
   3.5.7 Surface glycoproteins .......................................................................................... 97
   3.5.8 Viral Homologues of Chemokine Receptor and G Protein-Coupled Receptor
        Proteins .................................................................................................................... 97
   3.5.9 Immunomodulatory Genes ................................................................................. 98
   3.5.10 Anti-Apoptotic Genes ......................................................................................... 99
   3.5.11 Latency Genes .................................................................................................... 99
   3.5.12 Spliced Transcripts ............................................................................................ 100
   3.5.13 Missing genes ..................................................................................................... 100
3.5.14 Phylogenetic Analysis ........................................................................................................ 101
3.6 Discussion ................................................................................................................................ 102

Chapter 4 .......................................................................................................................................... 132

4 The Development of Cynomolgus Macaque Cytomegalovirus (CyCMV) as a Viral Vector for an HIV/SIV vaccine ........................................................................................................ 132

Overview .......................................................................................................................................... 133
4.1 Chapter Overview ....................................................................................................................... 133
4.2 Abstract ....................................................................................................................................... 134
4.3 Introduction ................................................................................................................................. 135
4.4 Materials and Methods ............................................................................................................. 137
  4.4.1 Cells and Virus .................................................................................................................... 137
  4.4.2 Generation of CyCMV-eGFP Recombinant Virus ............................................................ 137
  4.4.3 Cre/loxP Recombination Methodology .............................................................................. 138
  4.4.4 Generation of CyCMV-loxP ............................................................................................... 139
  4.4.5 Cloning CyCMV as a Bacterial Artificial Chromosome .................................................. 139
  4.4.6 Viral DNA Isolation ........................................................................................................... 140
  4.4.7 Electroporation .................................................................................................................. 140
  4.4.8 Polymerase Chain Reaction ............................................................................................... 140
4.5 Results ......................................................................................................................................... 141
  4.5.1 CyCMV-BAC Cloning Process ......................................................................................... 141
  4.5.2 Recombination Plasmids ................................................................................................... 143
  4.5.3 Generation of a CyCMV-expressing eGFP Recombinant Virus ...................................... 144
  4.5.4 Generation of CyCMV-loxP Intermediate ....................................................................... 146
  4.5.5 Cloning CyCMV as a Bacterial Artificial Chromosome .................................................. 147
4.6 Discussion ................................................................................................................................... 149
4.7 Acknowledgments ....................................................................................................................... 153
Chapter 5 ................................................................................................................................. 154

5 Discussion and Future Directions ......................................................................................... 154

5.1 Discussion ......................................................................................................................... 155

5.1.1 The Development of CyCMV as an HIV/SIV Vaccine Viral Vector .......................... 155

5.1.2 Rhesus Macaque Cytomegalovirus-Based Viral Vector .............................................. 158

5.1.3 Limitations and Safety Concerns to CMV-Based Vectors ........................................ 162

5.2 Future Directions ............................................................................................................ 164

5.2.1 CyCMV-eGFP-BAC Confirmation and Viral Reconstitution ................................. 164

5.2.2 Development of CyCMV-Based HIV/SIV Vaccine Candidates .............................. 165

5.2.3 Proposed CyCMV-SIV Vaccine Trial ........................................................................ 171

5.3 Conclusion ..................................................................................................................... 173

Appendices ............................................................................................................................. 175

References ............................................................................................................................... 185
List of Tables

Table 1-1: Comparison of VZV and CMV as HIV/SIV Vaccine Vectors................................. 44

Table 1-2 Natural versus Non-Natural SIV Hosts................................................................. 56

Table 3-1: CyCMV Gene Products....................................................................................... 106

Table 3-2: HCMV Genes Not Present in CyCMV ................................................................. 128
List of Figures

Figure 1-1 Global Distribution of HIV-1 Subtypes ................................................................. 2
Figure 1-2 HIV Structure ........................................................................................................ 4
Figure 1-3 HIV and SIV Viral Genomes ............................................................................... 5
Figure 1-4 HIV Replication Cycle ........................................................................................ 7
Figure 1-5 Clinical HIV-1 Infection ...................................................................................... 10
Figure 1-6 Impact of HAART on AIDS-Related Deaths ....................................................... 16
Figure 1-7 Hypothetical Course of HIV Infection in a Vaccinated vs. Unvaccinated Individual. 24
Figure 1-8: HIV Vaccine Strategies ...................................................................................... 27
Figure 1-9 Mucosal HIV/SIV Infection ................................................................................ 40
Figure 1-10 Timing of Effector Memory Responses and Impact on Viral Load ................. 42
Figure 1-11 Cytomegalovirus Structure .............................................................................. 48
Figure 1-12 Cytomegalovirus Genomic Organization .......................................................... 49
Figure 1-13: CMV Viral Replication Cycle ......................................................................... 50
Figure 1-14 CMV-Specific Memory T cell Phenotypes ........................................................ 54
Figure 1-15 Inoculation and Sampling Schedule ................................................................. 62
Figure 2-1 Electron Micrograph of CyCMV-Infected MRC-5 Cells. ................................. 72
Figure 2-2 Restriction Enzyme Digestion of CyCMV Genome ........................................ 74
Figure 2-3 Phylogenetic Analysis of DNA Polymerase and Glycoprotein B ..................... 75
Figure 2-4 Comparison of CyCMV Strain Differences in Glycoprotein B Sequences ....... 77
Figure 5-6 Viral Growth Kinetics of VZV-SIV Vaccine Constructs

170
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>Amino Acid</td>
</tr>
<tr>
<td>ADCC</td>
<td>Antibody-Dependent Cell-Mediated Cytotoxicity</td>
</tr>
<tr>
<td>AIDS</td>
<td>Acquired Immunodeficiency Syndrome</td>
</tr>
<tr>
<td>ART</td>
<td>Antiretroviral Therapy</td>
</tr>
<tr>
<td>ATCC</td>
<td>American Type Culture Collection</td>
</tr>
<tr>
<td>BAC</td>
<td>Bacterial Artificial Chromosome</td>
</tr>
<tr>
<td>BP</td>
<td>Base Pairs</td>
</tr>
<tr>
<td>CCR5</td>
<td>CC Chemokine Receptor 5</td>
</tr>
<tr>
<td>CD</td>
<td>Cluster of Differentiation</td>
</tr>
<tr>
<td>CM</td>
<td>Central Memory</td>
</tr>
<tr>
<td>CMV</td>
<td>Cytomegalovirus</td>
</tr>
<tr>
<td>CPE</td>
<td>Cytopathic Effect</td>
</tr>
<tr>
<td>CTL</td>
<td>Cytotoxic T-cell Lymphocyte</td>
</tr>
<tr>
<td>CXCR4</td>
<td>CXC Chemokine Receptor 4</td>
</tr>
<tr>
<td>CyCMV</td>
<td>Cynomolgus macaque Cytomegalovirus</td>
</tr>
<tr>
<td>DC</td>
<td>Dendritic Cell</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s Modified Eagle’s Medium</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribose Nucleic Acid</td>
</tr>
<tr>
<td>dsDNA</td>
<td>double stranded Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>dUTPase</td>
<td>deoxyuridine-triphosphatase</td>
</tr>
<tr>
<td>EBV</td>
<td>Epstein-Barr Virus</td>
</tr>
<tr>
<td>eGFP</td>
<td>Enhanced Green Fluorescent Protein</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme Linked Immunosorbent Assay</td>
</tr>
<tr>
<td>EM</td>
<td>Effector Memory</td>
</tr>
</tbody>
</table>
ENV – HIV envelop glycoprotein
FBS – Fetal Bovine Serum
FDC – Follicular Dendritic Cell
Gag – Group Specific Antigen
GALT – Gut-Associated Lymphoid Tissue
gB – Cytomegalovirus Glycoprotein B
gH – Cytomegalovirus Glycoprotein H
gL – Cytomegalovirus Glycoprotein L
gM – Cytomegalovirus Glycoprotein M
gN – Cytomegalovirus Glycoprotein N
gO - Cytomegalovirus Glycoprotein O
gp120 – HIV Envelope Glycoprotein 120
gp160 – HIV Envelope Glycoprotein 160
gp41 – HIV Envelope Glycoprotein 41
GPCR – G-Protein-Coupled Receptor
HAART – Highly Active Antiretroviral Therapy
HCMV- Human Cytomegalovirus
HHV-5 – Human Herpesvirus-5
HHV-6 – Human Herpesvirus-6
HHV-7 – Human Herpesvirus-7
HIV – Human Immunodeficiency Virus
HSV-1 – Herpes Simplex Virus Type 1
HSV-2 – Herpes Simplex Virus Type 2
ICA – Inhibitor of Caspase-8-induced Apoptosis
IE-1 – Cytomegalovirus Immediate-Early Protein-1
IE-2 – Cytomegalovirus Immediate-Early Protein-2
IFN-γ – Interferon-gamma
Ig – Immunoglobulin
IL – Interleukine
IN – Integrase
KBP – Kilo Base Pairs
kDa – kiloDalton
KSHV – Kaposi’s Sarcoma-associated Herpesvirus
LC – Langerhans Cell
LTR – Long Terminal Repeat
MA – Matrix
MDC – Monocyte-Derived Dendritic Cell
MGP – Membrane Glycoprotein
MHC – Major Histocompatibility Complex
MIA – Mitochondrial Inhibitor of Apoptosis
MOI – Multiplicity of Infection
MP – Membrane Protein
MSF – T – (Mellow Skin Fibroblast – Transduced) Cynomolgus macaque fibroblast cell line
MTMP – Multiple Transmembrane Protein
NC – Nucleocaspid
NEF – HIV Negative Factor
NGS – Next-Generation Sequencing
NK – Natural Killer
O/N – Overnight
ORF – Open Reading Frame
p24 – HIV Gag capsid protein
PBMC – Peripheral Blood Mononuclear Cell
PBS – Phospho-Buffered Saline
PCR – Polymerease Chain Reaction
PDC – Plasmacytoid Dendritic Cell
PFU – Plaque Forming Unit
PIC – Pre-Integration Complex
POL – HIV Polymerase gene
pp65 – Cytomegalovirus Polyprotein-65
PR – HIV Protease
qPCR – Quantitative Polymerase Chain Reaction
REV – HIV Regulator of Expression of Virion Proteins
RhCMV – Rhesus macaque Cytomegalovirus
RNA – Ribonucleic Acid
RPMI – Roswell Park Memorial Institute Media
RT – Reverse Transcription
SD – Standard Deviation
SEM – Standard Error of Mean
SFV – Simian Foamy Virus
SIV – Simian Immunodeficiency Virus
SSDB – Single-Stranded DNA-Binding Protein
TCID_{50} – Tissue Culture Infectious Dose
TCR – T-Cell Receptor
Telo – RF – Telomerase Rhesus macaque Fibroblast
TLR – Toll-like receptor
TNF-α – Tumor Necrosis Factor – alpha
TP – Tegument Protein
TRL – Cytomegalovirus Terminal Repeat Long
TRS – Cytomegalovirus Terminal Repeat Short
UL – Cytomegalovirus Unique Long Gene
US – Cytomegalovirus Unique Short Gene
VIF – HIV Viral Infectivity Factor
VPR – HIV Viral Protein R
VPU – HIV Viral Protein U
VPX – SIV Viral Protein X
VZV – Varicella Zoster Virus
WPI – Weeks Post-Infection
List of Appendices

List of Manuscripts Arising From This Thesis


IN PREPARATION


Chapter 1

1 Introduction

1.1 Human Immunodeficiency Virus

Human Immunodeficiency Virus (HIV) is a lentivirus belonging to the Retroviridae family and the causative infectious agent of Acquired Immunodeficiency Syndrome (AIDS) (101). In 1981, AIDS was first clinically discovered in homosexual men in the United States by the Centre for Disease Control (CDC) (123). HIV-1 was subsequently isolated in 1983 by two independent research groups, one in France led by Luc Montagnier and Françoise Barré-Sinoussi, and the other in the United States led by Robert Gallo; though only the French researchers, Montagnier and Barré-Sinoussi, received the 2008 Nobel Prize in Medicine in recognition of their work (53, 279). Shortly after the discovery of HIV-1, another genetically distinct but morphologically similar retrovirus was discovered in West African patients, termed HIV-2 (150). Humans do not represent the natural hosts of this retrovirus, and it is known that HIV-1 and HIV-2 were introduced into the human population by cross-species transmission of the simian version of HIV, known as Simian Immunodeficiency Virus (SIV), derived from chimpanzees (Pan troglodytes; SIVcpz) or sooty mangabeys (Cercocebus atys; SIVsm), respectively (745). HIV-1 is derived from zoonotic infection of humans with SIVcpz (286, 423), with the oldest documented case of HIV-1 infection occurring around 1959 in Kinshasa, the capital city of the Democratic Republic of the Congo (889, 916). Though this was not the first case of cross-species HIV-1 infection in humans, this was the first time the virus was able to disseminate throughout the human population as a catalyst to the HIV/AIDS pandemic (889). In the mid-20th century, Kinshasa, formerly known as Léopoldville, was a large commercial trade city that underwent urban expansion, with the rivers serving as a means of transportation in and out of the city, allowing the virus to spread between regions and countries (137, 744, 889). The original host of the HIV-2 strain was sooty mangabeys (sm), which transmitted its version of SIV (SIVsm) to humans giving rise to HIV-2, as well as experimental infection of rhesus macaques (Macaca mulatta) leading to SIVmac (194, 287, 359). HIV-1 is the dominant strain of the global pandemic, with HIV-2 primarily circulating in Western Africa (521), and therefore HIV-1 will be the primary focus from here on in.
HIV-1 is divided into lineages, which include four groups termed major (M), outlier (O), non-major and non-outlier (N), and putative (P), with group M contributing to the global pandemic spread (916), while N (758), O (201), P (642) are currently geographically localized and less prevalent. The M group can be further divided into nine subtypes A, B, C, D, F, G, H, J, K, and circulating recombinant forms (CRF) of these subtypes, with subtypes A-C being the most prevalent (817) (Figure 1-1). Clade C has the highest global prevalence, which accounts for 50% of worldwide HIV-1 infections and is primarily circulating in India and Africa, whereas in North America, the dominant circulating strain is clade B (348). The hypervariability of HIV-1 contributes to this extreme viral diversity, which exists both within and between a given clade, with 20% variation in the envelope protein sequence within a clade and up to 35% variation between clades (446) (reviewed in Section 1.2.1 Obstacles in the Development of an HIV Vaccine).

![Figure 1-1 Global Distribution of HIV-1 Subtypes](Reproduced with permission from (817), Copyright Massachusetts Medical Society)
In the 2012 UNAIDS Global Report, the worldwide prevalence of HIV was estimated to be 34 million (31.4-35.8 million), however the pandemic exhibits a disproportionate burden with 69% of the infected global population living in Sub-Saharan Africa (836). Though the number of people living with HIV has been increasing, in 2011 the estimated number of new HIV infections decreased by 20% since 2001 to 2.5 million (2.2-2.8 million) and most importantly, due to antiretroviral therapy (ART), the number of people dying from AIDS-related diseases declined by 24% since 2005 to 1.7 million (1.5-1.9 million) in 2011 (836). With regard to HIV epidemiology in Canada, in 2012, the Public Health Agency of Canada reported similar trends as the global report. In 2011, the estimated number of Canadians living with HIV increased to 71,300 with almost half (46.7%) of those individuals being men who have sex with men (MSM) (112). An estimated 3,175 (2,250 - 4,100) new HIV infections were reported in Canada in 2011, which, unlike the global rates of new infections, marks only a modest decrease compared to 2008 (112). With regard to the number of AIDS-related deaths in Canada, the data collected up to 2009 showed a continual decrease in the number of deaths since the late 1990’s, corresponding to the advent of ART (111). Unfortunately, the global projections for the HIV/AIDS pandemic in 2030 include an increase in the number of HIV/AIDS-related deaths to between 3.7 and 6.5 million according to the World Health Organization report, or between 9 and 10 million as per the United Nations report, with HIV/AIDS being the third leading cause of death worldwide (515, 835). These projections contain a fair amount of uncertainty due to the unknown effect of drug resistance on future rates of ART response, particularly in Africa. However, these striking increases clearly highlight the importance of developing an effective HIV vaccine.

1.1.1 HIV/SIV Structure and Genome

The virion structure and sequence of HIV shares a similar morphology, size, and genomic composition to other lentiviruses (307, 777). The HIV virion is a 100 – 120µm enveloped virion with a characteristic cone-shaped core that contains two copies of the single-stranded RNA (ssRNA) viral genome (Figure 1-2) (272). HIV is a ssRNA virus with long terminal repeats (LTRs) at either end of a ~9kb genome that encodes nine open reading frames (ORFs), which are transcribed into fifteen proteins (270). The nine ORFs of HIV-1/SIVcpz include Gag, Pol, Env, Nef, Tat, Rev, Vif, Vpr, and Vpu, whereas HIV-2/SIVmac differ by encoding Vpx in place of Vpu (Figure 1-3) (579). The gag gene encodes a polyprotein precursor, Pr55Gag that is cleaved into four core structural proteins including the matrix (MA), capsid (CA), nucleocapsid (NC),
and p6 proteins (270-272). Similarly, cleavage of the Gag-Pol precursor, Pr160\textsuperscript{Gag-Pol} gives rise to the polymerase (Pol) enzyme, which encodes three enzymes: protease (PR or PROT), reverse transcriptase (RT), and integrase (IN) (270, 272). The precursor for the envelope (Env) glycoproteins (gp), gp160 encodes the surface glycoproteins, gp120 (surface unit) and gp41 (transmembrane unit) that form the viral envelope membrane (270, 272). The functions of these proteins will be described in more detail in Section 1.1.2 (HIV Replication Cycle).

![Figure 1-2 HIV Structure][Reprinted from The Lancet (52), with permission from Elsevier]

The six accessory proteins, Nef, Tat, Rev, Vif, Vpr, and Vpu/Vpx, play a number of different roles (Figure 1-3). The negative factor (Nef) protein downregulates MHC class I on the surface of HIV-infected cells in an effort to evade the cytotoxic T lymphocyte (CTL) response (33, 163, 737). With regard to the initial stages of viral replication, the trans-activator of transcription (Tat) protein enhances the initiation rate of viral mRNA transcription by increasing the efficiency of the transcription complexes in the HIV-1 promoter region located in the 5’ LTR (33, 270, 890). The regulator of expression of virion (Rev) protein binds to the Rev response element (RRE) binding-site located in the Env coding region, and post-transcriptionally regulates the nuclear export of viral mRNA from the nucleus to the cytoplasm for translation and viral packaging (270, 775). The virus infectivity factor (Vif) promotes HIV infectivity and viral replication by inhibiting antiviral host restriction factors (255, 797). Viral protein R (Vpr) has multiple functions that promote the rate of viral replication including nuclear import of viral pre-
integration complexes, transcriptional transactivation, and cell cycle arrest (155, 272, 888). The two genes that distinguish between HIV-1 and HIV-2 are viral protein U (Vpu) and viral protein X (Vpx) (Figure 1-3). Vpu is unique to HIV-1/SIVcpz and plays a role in enhancing virion release, and CD4 degradation in the endoplasmic reticulum (156, 798, 879), whereas Vpx is unique to HIV-2/SIVmac and is involved in nuclear import of the pre-integration complexes (258, 409, 614).

**Figure 1-3 HIV and SIV Viral Genomes** [Reprinted from (579), with permission from Elsevier]

Genomic sequencing of these viruses revealed that SIVcpz is the most closely related simian counterpart to HIV-1, with the Pol protein sequence being the most conserved and Vpu being the most divergent protein between the two strains (374). The HIV-2 genome was more closely related to SIVmac with 75% nucleotide homology, compared to the HIV-1 genome, which is only 40% homologous at the nucleotide level with HIV-2 and SIVmac (126, 326, 853). SIV cross-species transmission has resulted in the evolution of host-adapted SIV strains that have been isolated and characterized from various non-human primates (NHPs) (268, 331, 359) (see Section 1.4 Non-Human Primate Models). In addition, there are virus challenge stocks derived from rhesus macaques experimentally infected with SIVsm and subsequently passaged several times in rhesus macaques prior to isolation, which are routinely used as challenge viruses in NHP studies of HIV/SIV pathogenesis, treatment, and vaccination (27, 425, 499). These virus challenge stocks include a cloned virus SIVmac239 (424), swarm viruses SIVmac251 (461, 577)
and SIVsmE660 (360), and chimeric SIV viruses encoding HIV proteins termed Simian-Human Immunodeficiency Virus (SHIV) (484, 748, 750). Both HIV and SIV strains infect the same target cells, utilize parallel receptors/co-receptors, and have an analogous replication cycle allowing for translational studies to be performed in macaques that inform HIV research in the areas of pathogenesis, treatment, and vaccine development.

1.1.2 HIV Replication Cycle

HIV attachment to the target cell membrane and the initiation of the entry process is mediated through the HIV envelope surface glycoprotein (gp120) that binds the CD4 receptor on activated and memory CD4+ T cells, macrophages, or dendritic cells (272). The binding of the variable 3 loop of gp120 to the CD4 receptor enables the virus to bind to the co-receptors C-chemokine receptor 5 (CCR5) or CXC chemokine receptor type 4 (CXCR4) located in close proximity to CD4 on the cell surface (Figure 1-4) (251, 372, 578, 701). Two viral phenotypes exist, R5 viruses that use CCR5 as a co-receptor and are macrophage-tropic, versus X4 viruses that use CXCR4 as a co-receptor and are T cell-tropic (13, 140, 213, 224, 251). During primary HIV-1 infection, R5 tropic viruses are the predominant viral phenotype that are transmitted and establish infection in activated memory CD4+ T cells, which express the CCR5 co-receptor (73, 917). This has been highly supported by the observed HIV-resistance in individuals with a homozygous 32 bp mutation in CCR5 (CCR5-Δ32), and a delay in progression to AIDS in individuals with a heterozygous CCR5-Δ32 mutation (205, 475, 725). X4 tropic viruses bind to CXCR4, which is predominantly expressed on naïve CD4+ T cells (73). X4 viral isolates typically appear during HIV-1 chronic infection, and are associated with increased disease progression due to mass CD4+ T cell depletion of multiple CD4+ T cell phenotypes (13, 170, 251, 686, 736).

Following the engagement of gp120 with the co-receptor, a conformational change occurs in the gp120 glycoprotein allowing the gp41 transmembrane subunit to be exposed, and to initiate fusion with the cellular membrane through the insertion of the gp41 fusion peptide, which forms a pore in the bilayer (92, 272-274, 729). The exact mechanism by which this occurs remains to be fully elucidated. During viral entry, Nef acts to disrupt the actin cytoskeleton of the cell allowing the viral core to be released into the cell where it disassembles, allowing the pre-integration complex (PIC), made up of genetic material and viral proteins, to be released into the
cytoplasm (Figure 1-4) (109, 248, 522). The reverse transcriptase (RT) enzyme encoded by the virus initiates the reverse transcription of the viral ssRNA into linear double-stranded cDNA (272), however HIV’s error-prone RT lacks a proofreading mechanism resulting in point-mutations and genetic hypervariability at a remarkable rate of $10^{-5}$ mutations/base pair/replication cycle (discussed further in Section 1.2.1 Obstacles in the Development of an HIV Vaccine) (500, 659). The HIV cDNA forms a PIC comprised of viral proteins (RT, IN, MA, NC, Vpr) and host proteins (248, 272, 550). Nuclear localization of the PIC is facilitated by microtubule networks that allow the viral DNA to migrate to the nuclear pore for nuclear import, mediated by HIV-1 protein Vpr or HIV-2 protein Vpx (258, 272, 522, 614). Once in the nucleus, the viral DNA either forms non-integrated one- or two-LTR circularized DNA that remains in the nucleus, or the viral integrase (IN) enzyme acts to integrate the linear viral DNA, with the help of cellular proteins, into the host cell chromosome (Figure 1-4) (102, 247, 272, 404, 754). At this juncture, the provirus becomes permanently integrated into the host genome and serves as a template for viral synthesis (272).

Figure 1-4 HIV Replication Cycle [Reprinted from (746), with copyright to Cold Spring Harbor Laboratory Press]
Transcription of the proviral DNA is initiated by host transcription factors, NF-kB, NFAT, SP1, and the HIV Tat protein, which bind to the 5’ LTR promoter of the HIV genome to enhance the efficiency of transcription and produce high levels of genomic RNA (33, 395, 430, 431, 574, 890). Viral RNA is synthesized by HIV’s RNA polymerase (Pol), which transcribes the proviral DNA into RNA (272, 429). Rev acts to export the unspliced (precursor polyproteins), partially spliced (Env, Vif, Vpr, Vpu), and multiply spliced (Tat, Rev, Nef) mRNA transcripts, as well as the viral RNA genome from the nucleus to the cytoplasm (250, 272, 493). With respect to immune evasion, the virus delays the synthesis of the viral proteins, until the end of the replication cycle in an effort to evade MHC presentation to CTLs (153). The spliced mRNA and viral genomic RNA are ready for packaging, while the unspliced mRNA is translated into structural Gag and Gag-Pol precursor polyproteins (272). The synthesized viral proteins in the form of precursor polyproteins traffic to the cell membrane where they assemble with the viral RNA genome for virion formation (Figure 1-4) (746, 805). Similarly, the envelope protein precursor gp160 is translated by the ribosomes located in the endoplasmic reticulum (ER) and shuttled via the Golgi complex, where gp160 is proteolytically cleaved into gp120 and gp41 subunits, which form a lipid bilayer at the cell surface that coats the viral particle (272, 520). The immature noninfectious virions then bud from the cell membrane where they only mature when the viral protease (PR) cleaves the Gag-Pol polyprotein generating an infectious retrovirus that can now infect new cells (438, 629, 805). The entire HIV-1 replication cycle takes approximately 1 to 2 days to complete (153).

1.1.3 HIV-1 Pathogenesis and Infection

1.1.3.1 Transmission

HIV-1 is most commonly spread by heterosexual transmission at mucosal surfaces, including the female and male genital tracts, and the gastrointestinal (GI) tract, with transmission occurring in the form of cell-free or cell-associated virus via vaginal secretions, semen, rectal sections, or blood (705). Transmission can occur at different sites including the vagina, ectocervix, or endocervix in the female genital tract, the inner foreskin or penile urethra in the male genital tract, and the rectum or upper GI tract. The risk of acquisition (range of transmission probability per exposure event) varies depending on the site of infection and the mode of transmission (362). The most common site of HIV-1 infection worldwide is the genital tract, which has a relatively low risk per coital act with a higher risk of infection in the females
(1 in 200 to 1 in 2000) than the males (1 in 700 to 1 in 3000) (78, 320, 362, 610, 667). Though the risk of genital HIV acquisition in men is relatively low, there are other factors, such as co-infections with sexually transmitted infections (STIs), stage of HIV infection (Figure 1-5) (868), or penile circumcision, that can significantly impact the risk of infection (78, 656, 871). A number of different STIs can increase the infectiousness of HIV-1 and/or the susceptibility to infection in both men and women by increasing viral shedding leading to a higher risk of transmission, or altering the mucosal barrier and immune milieu in the genital tract making the individual more susceptible to HIV acquisition (257, 280, 700). Male circumcision has been shown to reduce the risk of HIV-1 infection and represents a strong public health preventative measure to reduce HIV infection in men during heterosexual intercourse (37, 44, 319, 821).

Other routes of transmission include sexual transmission via anal receptive intercourse, which has a higher risk of infection in the rectum (1 in 20 to 1 in 300) (41, 362). HIV-1 can also be vertically transmitted from mother-to-child by maternal blood or genital secretions intrauterine or during delivery, or subsequently transmitted in the intestinal tract via breast milk transmission during feeding (458, 586). However, the heightened risk of mother-to-child transmission (1 in 5 to 1 in 10) can be reduced with prophylaxis ART administered throughout the pregnancy, delivery, and breastfeeding stages (169, 178, 362, 458, 586). Finally, HIV-1 can also spread via the bloodstream through blood transfusions, Injection Drug Users (IDUs), or sharps pricks in healthcare workers. This mode of transmission confers the highest risk (95 in 100 to 1 in 150) as the bloodstream lacks a protective mucosal barrier and is thus a direct route for peripheral dissemination of the virus (362).

1.1.3.2 Pathogenesis and Infection

1.1.3.2.1 Acute Infection

Primary HIV-1 infection is characterized by a rapid peak in HIV-1 plasma viral load paired with a decline in CD4+ T cells, followed only after by an increase in HIV-specific CD8+ CTLs (Figure 1-5). During acute HIV-1 infection, symptoms arise 2-4 weeks post-transmission, often last 3-4 weeks and resemble influenza-like or mononucleosis-like symptoms, which can include fever, headache, sore throat, swollen lymph nodes, joint or muscle pain, malaise, rash, and weight loss (187, 189, 346, 828). In addition to symptomatic diagnosis, acute HIV-1 infection can be clinically diagnosed by serologic assays specific for HIV-1 Gag proteins or
HIV-specific antibodies, such as the enzyme-linked immunosorbent assay (ELISA) and Western blot, or by virologic assays in which HIV-1 RNA can be detected and quantified in the blood (copies/ml) by quantitative polymerase chain reaction (qPCR) assay (187, 189). Furthermore, enumeration of T cells (cells/mm³) in the peripheral blood can demonstrate a decrease in absolute CD4⁺ T cell count and an increase in HIV-specific CD8⁺ T cells (as shown in Figure 1-5).

**Figure 1-5 Clinical HIV-1 Infection** [Reprinted from The Lancet (759), with permission from Elsevier]

The vast majority of the information regarding the early stages of HIV-1 infection stems from seminal research done in non-human primates infected with SIV, which enables researchers to temporally monitor SIV infection at various mucosal sites (2, 328, 329, 369, 549, 783). During
primary HIV-1 infection, the transmission medium (i.e. mucosal secretions, semen) contains diverse viral variants, however a genetic bottleneck phenomenon occurs during mucosal transmission resulting in only a single founder viral variant passing the mucosal barrier in 80% of HIV-1 infections (311, 327, 422, 723, 917) (Figure 1-5). Alternatively, a study examining female sex workers that were infected with multiple viral variants found factors such as sexually transmitted infections and hormonal contraceptives to be associated with increased viral diversity during primary HIV-1 transmission (720).

Upon passing the mucosal barrier, HIV-1 envelope glycoprotein gp120 binds to the CD4 receptor, CCR5/CXCR4 co-receptors, and a specific Langerin receptor uniquely expressed on intraepithelial Langerhans cells (LCs), resulting in viral infection in LCs, which facilitates the transport of the virus to the submucosa (246, 369, 548, 778, 907). In the submucosa HIV-1 infection is initiated by binding of the HIV-1 envelope glycoprotein gp120 to the HIV receptor CD4, and co-receptor CCR5 expressed on dendritic cells, macrophages, or T cells (190, 433, 653). Dendritic cells express a C-type lectin DC-SIGN, which captures HIV allowing for trans infection and transportation of the virus to resting and activated CD4+ T cells, which are the preferential target cell population of HIV and SIV (292, 363, 912). Dendritic cells and macrophages residing in the lamina propria can also be productively infected with HIV-1 (289, 437, 565, 624, 782, 832). The infected CD4+ T cells are able to carry the virus further into the tissue where it begins local expansion in the regional lymphoid tissues (401). This local expansion occurs during the eclipse phase of infection, which lasts 1 to 2 weeks post-infection and is a period in which the virus is not yet detectable in the plasma (153, 528).

1.1.3.2.1.1 Systemic Infection

Following sufficient local viral replication and virion production, the virus disseminates to the periphery via the draining lymph nodes, which is marked by seroconversion and the initiation of the acute phase of infection that typically lasts 2-4 weeks (153). At this stage, viral entry into the bloodstream is not met with a primed HIV-specific immune response, which facilitates high levels of viral replication in activated peripheral CD4+ T cells and a burst of plasma viremia that establishes the peak viral load (Figure 1-5) (149, 759). An inverse relationship exists between the level of viremia and CD4+ T cell count, in which the burst in viremia is mirrored by a major decline in CD4+ T helper cell population causing substantial
cellular and humoral immune dysfunction (312, 452, 511, 759). In the later stage of acute infection, plasma viral load is rapidly reduced following the establishment of the HIV-specific CD8\(^+\) CTL response (85, 188, 442). In addition, HIV-specific antibodies respond to viral proteins, and although the early founder viruses are neutralization-sensitive, neutralizing antibodies do not develop until approximately 12 weeks post-infection (reviewed in Section 1.1.4.2 Adaptive Immune Response) (177, 216, 559). This enhanced immune pressure promotes viral escape mutations to evade both the cellular and antibody-mediated immune response, contributing to neutralization-resistant viruses and a progressive increase in viral diversity throughout the course of HIV infection (Figure 1-5) (86, 216, 687, 869). The immune response to HIV infection is discussed in detail in Section 1.1.4 (Immune Response to HIV Infection). The partial control of plasma viremia by the antiviral activity of the CTL response enables a transient rebound in the CD4\(^+\) T cell population and establishment of the viral load set-point (Figure 1-5). Viral set-point varies between individuals and is established following acute infection when the viral load reaches a relatively steady state, which occurs during the chronic phase of infection (153).

1.1.3.2.1.2 Lymphoid Tissue Infection

Once circulating at high levels in the periphery, the virus rapidly disseminates to various tissue sites, including the gut-associated lymphoid tissues (GALT), distal lymph nodes, spleen, and the brain, where it establishes a persistent viral reservoir (401). Upon entry into the secondary lymphoid organs, virions are trapped by the follicular dendritic cell (FDC) network, which creates immune complexes that are subsequently converted to highly infectious forms and are able to infect CD4\(^+\) T cell and macrophages passing through the lymphoid follicles (345, 616, 730). Lymphoid tissues, particularly the GALT, are rich in target cells, which allows for massive infection of the effector memory CD4\(^+\) T cell population and considerable destruction of the lymphoid architecture (516, 619). The GALT target cells, specifically Th17 and Th22 CD4\(^+\) T cell subsets, are highly susceptible to infection as they have increased expression of a gut-homing marker, \(\alpha 4\beta 7\), which is an integrin involved in trafficking T cells to mucosal sites and binds HIV gp120 to facilitate infection through the CD4 and CCR5 receptors that reside in close proximity to \(\alpha 4\beta 7\) (32, 147, 333, 400, 849, 852). This is mirrored by a major depletion in CD4\(^+\) T cells, which is most pronounced in the gastrointestinal tract and results in immune dysfunction and mass destruction of the GALT, as observed in HIV-infected patients and following SIV infection.
in NHP models (Figure 1-5) (98, 324, 462, 471, 516, 531, 842). Furthermore, the T cells in the gastrointestinal tract are predominantly memory CD4+ CCR5+ CXCR4+ T cells making them highly susceptible to HIV-1 infection and viral replication, resulting in pronounced mucosal CD4+ T cell depletion (533, 651).

1.1.3.2.2 Chronic Infection

The chronic stage of HIV-1 infection lasts for approximately 8 to 10 years in the absence of antiretroviral therapy (ART), during which time patients are routinely monitored for CD4+ T cell count and HIV-1 viral load to determine if ART should be initiated or if already on treatment, to monitor for evidence of drug resistance that may require alterations in the drug regimen. Though typically asymptomatic, the immune dysfunction causes some individuals to experience fatigue, weight loss, night sweats, oral lesions (hairy leukoplakia), oral and vaginal candidiasis, and symptoms associated with herpesvirus reactivation (446). In addition, muscle wasting, dermatological conditions, neurological complications, and multiple organ dysfunction, including cardiovascular disease, can be observed during chronic HIV infection (179, 446, 576, 760, 865).

The partial control of infection by the antiviral immune response reduces the peak viral load to a steady state set-point during chronic HIV-1 infection (Figure 1-5). Set-point plasma viral load is the best predictor of HIV-1 disease progression to AIDS and AIDS-related neurologic dysfunction, followed after by CD4+ T cell count as a predictor (136, 536, 537). The variation in viral load set-point between individuals can be attributed to the magnitude and quality of the immune response (reviewed in more detail in Section 1.1.4 Immune Response to HIV Infection) (436, 617, 690), and a number of different genetic factors, such as HLA alleles (HLA B*57 and HLA B*27) (20, 43, 349, 547, 580) or a heterozygous CCR5-Δ32 mutation (previously discussed in Section 1.1.2 HIV Replication Cycle) (555). Long-term non-progressors (LTNPs) and elite controllers are therapy-naïve HIV-infected individuals that naturally have low or undetectable viral load, respectively, maintain healthy CD4+ T cell counts, have a high quantity and quality of functional anti-HIV CD8+ T cells, and most importantly do not progress to AIDS or progress very slowly over time (68, 117, 207, 545, 546, 555, 620, 690).

At the beginning of the chronic phase of infection, there is minimal change in viral load and CD4+ T cell counts, however over time, a slow and quasi-steady increase in viral load is
paired with a progressive and gradual decrease in the CD4+ T cell population (Figure 1-5) (153, 759). Once in the lymphoid tissues, the virus rapidly establishes a persistent tissue viral reservoir that is maintained throughout the course of infection and even during prolonged ART, making it difficult to eradicate the virus from the host (144, 555). When HIV-1-infected activated CD4+ T cells revert to a long-lived resting memory CD4+ T cell phenotype, a latent reservoir is formed in the lymphoid tissues (141, 142). Latently infected resting CD4+ T cell reservoirs harbor integrated HIV-1 provirus, are transcriptionally silent, and have a slow rate of decay, which presents significant challenge for HIV-1 eradication as they are not eliminated with ART (145, 253, 755, 756, 887, 911). Though ART patients may have undetectable plasma viral load, these latently infected viral reservoirs harbor replication-competent proviral DNA that continues to replicate at low levels, which replenishes the reservoir and mediates viral relapse when ART is interrupted (143, 197, 253, 278, 743, 911). Upon reactivation of latently infected resting CD4+ T cells, the virus becomes transcriptionally active allowing for HIV-1 virions to be produced and infect activated CD4+ T cells, which further promotes replenishment of the persistent latent viral reservoir and enhances the complexity of viral eradication (143).

The depletion of the CD4+ T cell compartment in the GALT is maintained during chronic HIV-1 infection (Figure 1-5), and does not mimic the immune reconstitution observed in the peripheral blood observed following the initiation of ART (144, 364, 532). HIV-1 preferentially infects and causes profound depletion in the Th17 and Th22 CD4+ T cell subsets, which defend against microbes and play a role in maintaining the integrity of the gastrointestinal tract mucosal barrier, thereby preventing microbial translocation from the GI tract to the periphery (95, 191, 428, 527, 657, 669). The depletion of these cells leads to enhanced gut permeability allowing for microbial translocation, which in turn leads to increased and sustained systemic immune activation, a hallmark of chronic HIV-1 infection (96, 97). Chronic immune activation is likely driven by multiple factors that promotes widespread production of proinflammatory factors and increased viral replication in activated T cells (555). This cellular hyperactivity leads to persistent high T cell, B cell, and NK cell turnover, which further contributes to increased immune dysfunction marked by cellular exhaustion, senescence, and limited renewal capacity (200, 459, 519, 554, 555). In this regard, chronic systemic immune activation has a poor prognosis for survival and is associated with a more rapid progression to AIDS (298, 344).
The onset of Acquired Immunodeficiency Syndrome (AIDS) occurs following a pronounced depletion of both naïve and memory CD4$^+$ T cells and is diagnosed when the CD4$^+$ T cell count declines below 200 cells/mm$^3$ (HIV-negative CD4$^+$ count range: approx. 600-1600 cells/mm$^3$), which is matched by a rapid increase in HIV plasma viral load (Figure 1-5) (446). This immunosuppression allows infections to surface, which can include oral candidiasis, pneumococcus, tuberculosis, and the reactivation of latent herpesviruses, such as herpes simplex virus, varicella zoster virus, or cytomegalovirus (446). In addition, reactivation of oncogenic herpesviruses, such as Epstein-Barr virus or Kaposi’s Sarcoma herpesvirus, can cause AIDS-related malignancies (446). As the CD4$^+$ T cell count continues to drop below 200 cells/mm$^3$, several major opportunistic infections are classically observed including pneumonia (*Pneumocystis*), *Mycobacterium avium complex*, cryptococcosis, and cytomegalovirus retinitis caused by cytomegalovirus reactivation (512, 612). Once the CD4$^+$ T cell counts drop to below 50 cells/mm$^3$, these opportunistic infections are typically the AIDS defining illnesses that are fatal (446). However, since the advent and initiation of Highly Active Antiretroviral Therapy (HAART), the morbidity and mortality associated with HIV infection has greatly decreased, with the number of AIDS-related deaths being inversely related to the number of patients on HAART (Figure 1-6) (611, 612). Antiretroviral therapy has been one of the greatest medical discoveries of our time with great success in treating HIV-1 infection, delaying the progression to AIDS, reducing the number of AIDS-related deaths, and even decreasing the risk of transmission (160, 221, 611). HAART is effective at rapidly decreasing viral load in the form of cell-free virus or virus in cells with a short half-life, however it fails to eliminate the viral reservoir that is maintained in the long-lived resting CD4$^+$ T cell population (145, 253, 631, 887). Even after 7 years on therapy, HIV-1 viremia is still detectable in viral suppressed patients (613). To circumvent this, researchers are examining a number of novel therapeutic methods to eradicate the latent viral reservoirs in patients on HAART in an attempt to cure these individuals of HIV-1 infection (206).
1.1.4 Immune Response to HIV Infection

The host immune system mounts a robust response to HIV infection that has been extensively studied since the discovery of the virus; however the immune correlates of protection remain to be elucidated. The virus employs a number of immune evasion strategies that impairs the immune response resulting in only transient control of viral replication and HIV pathogenesis (in the absence of ART), thus preventing viral eradication. The first intrinsic antiviral defense to HIV infection is derived from host innate restriction factors that control HIV/SIV species tropism by blocking cross-species infection and restricting early post-entry HIV-1 replication (184, 357, 366, 749, 799). These germ-line encoded restriction factors primarily include APOBEC3G (apolipoprotein B mRNA editing enzyme catalytic polypeptide-like 3G) (139, 184, 497, 747), TRIM5a (tripartite motif 5a protein) (801), and tetherin (582, 583). HIV is able to overcome these host defense mechanisms by encoding viral proteins that function to inhibit or degrade the restriction factors, such as Vif mediated degradation of APOBEC3G (502, 534, 794, 904) or Vpu downregulation of tetherin (583, 840). Failing the intrinsic host cell restriction defense, the next HIV antiviral response comes from the innate immune response.
1.1.4.1 Innate Immune Response

The innate immune response serves to prevent or limit HIV infection and viral replication at the portal of entry in an effort to prevent viral dissemination to the lymphoid tissues. The early innate immune response begins at the mucosal barrier to prevent HIV entry and viral replication. One of the first cell types infected by HIV-1 following sexual transmission are Langerhans cells (LCs) (previously reviewed in Section 1.1.3.2.1 Acute Infection), which express a unique HIV-1 gp120 receptor, Langerin, that acts as a defense mechanism by facilitating viral internalization and degradation in cytoplasmic organelles, known as Birbeck granules (204). The innate immune cells involved in combating HIV infection includes macrophages, dendritic cells (DCs), and natural killer (NK) cells. Macrophages secrete a number of different cytokines (i.e. IL-1β, IL-6, IL-12, TNFα, IFNγ) and antiviral chemokines (RANTES, MIP-1α, MIP-1β) that function to block HIV entry or inhibit various stages of viral replication and virion assembly (115, 180, 242, 556). Monocyte-derived DCs (mDCs) sense the virus via toll-like receptors (TLRs), which trigger the release of type 1 interferon, pro-inflammatory molecules IL-6, TNFα, MIP-1α, and IL-12, IL-15, IL-18 that mediate NK cell activation and NK priming of the antiviral adaptive immune response (252, 486). Similarly, DCs act as antigen-presenting cells (APCs) to prime the adaptive immune response, specifically HIV-specific antiviral CTLs (161, 714, 908). However, internalization of intact virions inhibits the activation of mDCs, which is required to initiate the adaptive immune response (263). Viral ssRNA and unmethylated CpG DNA is also detected by TLR7 and TLR9, respectively, expressed on plasmacytoid dendritic cells (pDCs), which triggers the release of type 1 interferons (IFNα) that act to inhibit viral replication and activate NK cells (57, 252, 463, 496, 535, 681, 696). The NK cell response plays a crucial antiviral role during early infection by hindering HIV infection indirectly through the release antiviral cytokines (IFNγ and TNFα) and chemokines (MIP-1α, MIP-1β, RANTES), and also by directly lysing HIV-infected cells with cytotoxic mediators (perforin and granzyme A), however in the later stages of infection, HIV viral load is inversely correlated with NK cytolytic function (18, 249, 441, 605, 654). An additional subset of NK cells are involved in antibody-dependent cell-mediated cytotoxicity (ADCC) and exhibit increased activity in HIV-infected individuals (18), which will be discussed in detail in Section 1.1.4.2 (Adaptive Immune Response). The importance of the NK cell response is reflected in the HIV encoded proteins that are directed at evading NK cell-mediated cytotoxicity. NK cells characteristically kill cells that do not express
MHC class I on their cell surface, and although the HIV-1 viral protein Nef downregulates MHC class I to evade the CTL response, it selectively downregulates HLA-A and -B, known to present CTL-restricted epitopes, but retains HLA-C and –E, which are the ligands for the NK cell inhibitory receptors, thereby preventing NK lysis of virally infected cells (79, 157, 163).

The cytokine and chemokine response acts as a double-edged sword recruiting dendritic cells, macrophages, and activated CD4⁺ T cells to the site of infection, giving HIV increased access to the accumulating target cells (816). During the eclipse phase of HIV infection, pro-inflammatory cytokines are released at the mucosal sites creating a cytokine storm leading to immune activation, the recruitment of HIV target cells, and increased viral replication (2, 321, 784). Though HIV-1 infection leads to a depletion of a large number of mDCs and pDCs, the remaining DCs exhibit a dysfunctional immune response resulting in hyperimmune activation, impaired functional capacity, and type 1 IFN induced-apoptosis of CD4⁺ T cells contributing to enhanced HIV immunopathogenesis (219, 220, 350, 351, 609, 715, 779). Furthermore, HIV-induced TLR signaling in pDCs leads to the release of chemoattractants (MIP-1α, MIP-3α, MIP-1β, RANTES) that recruit CD4⁺ T cells to the site of infection fueling viral dissemination in the target cell population (463). In addition, HIV activation of pDCs upregulates the expression of indoleamine 2,3 dioxygenase (IDO), which facilitates naïve CD4⁺ T cell differentiation into regulatory T cells (Tregs) (495). Studies in non-human primate models that do not progress to AIDS (i.e. natural hosts), and also in HIV-infected children, have demonstrated that an increased early systemic anti-inflammatory cytokine response (TGF-β, IL-10) might be protective by inducing a Treg response, which regulates immune activation thus delaying disease progression (439, 440, 751, 824). Conversely, during chronic HIV infection, microbial translocation leads to upregulation of IL-10 production by monocytes in an attempt to inhibit chronic immune activation, however this dampens the antiviral immune response and contributes to disease progression (330, 445, 721, 802). It is evident that the innate immune response is not robust enough to inhibit HIV pathogenesis independently and requires further support from the adaptive immune response.

1.1.4.2 Adaptive Immune Response

The adaptive immune response involves priming from the innate immune system resulting in a delay in the appearance of HIV-specific antibodies and the initiation of the antiviral
T cell response. The humoral immune response to HIV is multifaceted, including monoclonal antibodies, as well as antibodies involved in viral neutralization, antibody-dependent cell-mediated cytotoxicity (ADCC), and antibody-dependent cellular viral inhibition (ADCVI). Monoclonal HIV Env-specific IgM and IgG antibodies are the first humoral responders that form immune complexes, with anti-gp41 antibodies appearing ~3 weeks post-infection followed two weeks later by anti-gp120 antibodies (1, 528, 829). Furthermore, anti-gp41 IgA secreted at the mucosa are detected in low levels very early during acute HIV-1 infection (540, 902). In addition to the indirect effect of CD4+ T helper cell depletion on B cell maturation, HIV-1 infection directly disrupts the antibody response through increased B cell apoptosis, as well as Nef-mediated B cell hyperactivation and suppression of class-switch recombination (491, 492, 567, 666, 808). Therefore, these initial antibody responses are ineffective at reducing HIV-1 plasma viremia and do not exert strong immune pressure on the virus, which is evident by a lack of viral escape mutations to this type of humoral response (422, 829). Non-neutralizing antibodies demonstrate antiviral functions (ADCC, ADCVI) that aid the CTL response in controlling acute HIV viremia by targeting virus-infected cells and cell-free virions (168, 266), which have been shown to be associated with reduced viremia and delayed disease progression in non-human primate studies (48, 260, 304, 354, 893). ADCC involves HIV-specific antibody binding to HIV Env molecules expressed on the cell surface of infected cells and simultaneous engagement of the Fc portion of the antibody to the Fc receptor expressed on NK cells to mediate NK cell cytotoxicity and the elimination of the infected cells (266, 353, 796). Similarly, ADCVI mediates ADCC in infected cells and in addition, HIV-specific IgG antibodies are also involved in the elimination of cell-free virus via Fc-mediated phagocytosis by NK cells, and NK cell-induced antiviral cytokine and β-chemokine inhibition of virions by blocking HIV entry (265).

The HIV virion contains neutralization-sensitive antigens within the Env subunits, gp41 and gp120, that enable neutralizing antibodies to inhibit the virus (872). Neutralizing antibodies use the Fc portion to block HIV-1 Env trimers, which prevents cell-free virus from binding to cellular receptors and co-receptors on target cells, thus preventing viral entry and cell-to-cell spread following infection (353). Passive transfer of HIV-specific neutralizing antibodies in non-human primates and humans have demonstrated the effectiveness of these antibodies in preventing mucosal infection or viral rebound after therapy termination, respectively (38, 243, 507, 509, 558, 831, 843, 894). Though highly effective at preventing HIV-1 acquisition,
naturally developed neutralizing antibodies have limited breadth, are present in low titers, and generally only appear 3 months after the onset of symptoms (1, 317, 869), which is too late to prevent or control acute HIV infection and widespread viral dissemination. Furthermore, only 20% of HIV-infected individuals develop broadly cross-reactive neutralizing antibodies with an increased breadth that target both gp41 and gp120, and even more rare are individuals with antibodies capable of neutralizing the CD4 binding site of gp120, however these antibodies only develop after more than two years post-infection (222, 316, 448, 488, 528). Furthermore, once neutralizing antibodies do appear and begin exerting immune pressure on the virus, the virus rapidly acquires escape mutations and employs structural evasion mechanisms leading to neutralization-resistant virions (reviewed in more detail in Section 1.2.1 Obstacles in the Development of an HIV Vaccine) (687, 869).

The cellular arm of the adaptive immune response plays a critical role in controlling HIV replication and disease progression. CD4\(^+\) T cells are the most impacted by HIV, yet still manage to exert an effective antiviral response to the virus. HIV-specific CD4\(^+\) T follicular helper cells provide pivotal help to follicular B cells in the germinal centers, facilitating B cell affinity maturation and the production of virus-specific antibodies (632). Moreover, HIV-1-specific CD4\(^+\) T cells exert a direct effector response through the secretion of antiviral cytokines (IFN\(\gamma\)), cytotoxic mediators (perforin, granzymes), and \(\beta\)-chemokines (RANTES, MIP-1\(\alpha\), MIP-1\(\beta\)), which in some cases have been shown to be inversely correlated with HIV-1 viral load (67, 203, 301, 637, 699, 776). Furthermore, CD4\(^+\) T cell subsets, including Th17 and Th22 cells, exhibit protective mucosal immunity through the release of IL-17 and IL-22 that preserve the homeostasis and integrity of the mucosal barrier, thus reducing microbial translocation and systemic immune activation (95, 428, 669, 800). Lastly, HIV-specific CD4\(^+\) helper T cells play a critical role in initiating the highly effective CD8\(^+\) T cell response (402).

Following the peak in viral load, the HIV-specific CD8\(^+\) T cell response arises at its highest level to control viral replication and reduce viremia to a set-point level (Figure 1-5) (85, 309, 442, 618, 881). HIV-specific CD8\(^+\) T cell recognition of viral peptides presented by MHC class I mediates effector cytotoxic T lymphocyte (CTL) activity in HIV-infected individuals (365, 857). The initial CTL responses that are detectable during the first 4 weeks following seroconversion have a limited epitope breadth and are specific for Env and Nef, however the breadth of CTL recognition increases to include Gag and Pol epitopes, followed thereafter by a
broad response that recognizes the remaining viral proteins by 6 months post-infection (5, 85, 116, 442, 450). With respect to the frequency of responses, HIV Gag and Nef are the most frequently recognized proteins by the CTL response (5, 467, 834). The direct effect of the CTL response on HIV is mediated by both noncytolytic and cytolytic mechanisms through the release of β-chemokines (RANTES, MIP-1α, MIP-1β), cytolytic molecules (perforin, granzymes), and antiviral cytokines (IFNγ, TNFα), which block HIV entry, lyse infected cells, and suppress viral replication, respectively (152, 490, 546, 858, 898, 899). Increased HIV-specific CTL activity is inversely correlated with viral load, and is associated with more steady CD4+ T cell counts overtime, which strongly showcases the role of CTLs in HIV viral control and the maintenance of viral set-point during chronic infection (85, 86, 323, 573, 602). Once AIDS develops, the continued decline in CD4+ T helper cells paired with increased viral diversity and CD8+ T cell exhaustion, makes it difficult for the CD8+ T cell response to control viral replication, which is mirrored by an increase in viremia and the onset of opportunistic infections (Figure 1-5) (365, 394, 436).

The importance of the CTL response in controlling viral replication and disease progression has been highlighted in a number of different circumstances. HIV-exposed uninfected women or exposed but uninfected infants, exhibit an HIV-specific CD8+ T cell response suggesting protective CTL-mediated immunity, and furthermore in women that did become infected, the pre-existing HIV-specific CTL responses were associated with a lower set-point viral load and a dysfunctional antiviral T cell response (414, 415, 702-704). Moreover, HIV-infected long-term non-progressors (LTNPs) that have low viral load and do not progress to AIDS, maintain high quantity and quality of polyfunctional anti-HIV CD8+ T cells (68, 117, 545, 546, 690). Similarly, the association between HLA class I alleles (HLA B*57 and HLA B*27) and a delayed progression to AIDS emphasizes the role of the CTL response in controlling viral replication (20, 43, 349, 547, 580). The impact of the CTL response on viral control was further confirmed when CD8+ T cells were experimentally depleted in SIV-infected non-human primates and a rapid and dramatic increase in plasma viral load was observed (386, 514, 732). Furthermore, the first viral escape mutations observed during the early stages of acute infection are in the CTL epitopes and the selective pressure of the CTL response is reflected by viral escape mutations throughout the course of infection (14, 66, 309, 597, 633). The earliest CD8+ T cell escape mutations occur in Env and Nef during early acute infection, followed thereafter by
CTL-selected viral escape mutations in Gag and Pol in the later stages of acute infection (309, 864). The increase in viral escape mutations paired with increased HIV-specific CD8\(^+\) T cell exhaustion hinders the CTL response, which is then unable to control viral replication leading to increased viremia and disease progression (199, 291). In this regard, it is evident that the host immune response is incapable of efficiently clearing HIV infection and therapeutic interventions are required to combat the virus and reduce the viral burden imposed on the host immune system. The HIV research field is investing considerable efforts on cure research by examining therapeutic methods of eradicating HIV. This is in addition to the longstanding focus of HIV vaccine researchers striving to develop a preventative and/or therapeutic vaccine, which will be the focus moving forward in this thesis.

1.2 HIV/SIV Vaccine Development

Despite major advances in antiretroviral therapy and the research on novel therapeutic methods, there is still no cure and the HIV/AIDS pandemic continues to spread. In fact, for every person starting ART, two more are newly infected with HIV (837). A number of other therapeutic strategies are being explored to treat HIV infection in individuals on ART. The highly publicized strategy that has gained significant media attention is the possibility of eradicating HIV via stem cell transplantation. The best-documented case of this was the “Berlin Patient” (15, 376). Timothy Brown was an HIV-1-positive male who had acute myeloid leukemia (AML) and received a bone marrow transplant from a CCR5-Δ32 donor. A Δ32 mutation in the CCR5 allele, the co-receptor of HIV CCR5-tropic viruses, causes a 32 base pair deletion in the CCR5 gene, which has been shown to make CD4\(^+\) T cells resistant to R5-tropic HIV infection (205, 475, 725). Therefore, the transplanted CD4\(^+\) T cells were resistant to HIV infection, thus clearing Timothy Brown of HIV-1, which to-date has been undetectable for over 3.5 years post-CCR5-Δ32 CD4\(^+\) T cell transplant (15, 376). The difficulty with repeating this method of eradication in other HIV-positive patients is the life threatening risks involved with bone marrow transplants, which include pre-transplant immunosuppression, total body irradiation, and graft-versus-host disease. Furthermore, another restriction is the limited availability of HLA-matched CCR5-Δ32 donors, which occurs in Caucasians with a frequency of less than 10% depending on the geographic region (487, 506, 919). These risks and challenges make bone marrow transplantations an unlikely therapeutic option for eradicating HIV.
Thus, an HIV vaccine remains the best method for controlling the HIV pandemic and vaccine research needs to be a global priority. Of course the ultimate goal for an HIV vaccine would be to elicit sterilizing immunity, thus completely preventing or aborting HIV infection. However, a vaccine that induces partial protection, in which individuals do become infected but the vaccine reduces pathogenesis by lowering viral load and preserving the CD4+ T cell population, will also be advantageous. Studies have shown that a vaccine capable of lowering HIV viral load by 0.5 \( \log_{10} \) copies/ml or greater would not only have clinical benefits for the disease course of the patient, but it would also have important implications for preventing secondary HIV transmission (320, 325, 667).

Vaccines are designed to induce a rapid memory recall response allowing for primed lymphocytes to block and/or eliminate the pathogen at the site of exposure. The goal for a successful HIV vaccine is to be able to stop HIV replication as early as possible, thus reducing the window of vulnerability (Figure 1-7) (393). Early immune intervention is imperative for controlling infection, thus a vaccine should be designed to inhibit HIV at the portal of entry and within the first week of infection, before the virus leaves the site of primary infection and undergoes a peak burst of viremia (329, 393). After this point, the virus irreversibly disseminates to the periphery where it is able to form a large HIV reservoir leaving the immune system trying to catch-up with the virus (634). Furthermore, another temporal challenge for the vaccine response is once the virus establishes systemic infection it will become latent in resting CD4+ T cells, where it will be hidden from a vaccine-elicited immune response (72, 393).
The limited insight into the immune correlates of protection has made HIV vaccine design challenging, however current clinical efficacy trials are beginning to shed light on areas of vulnerability within the virus that may confer protection (203, 343, 410, 466, 626). Thus, a vaccine will need to exploit and target HIV’s areas of immune vulnerability, which are becoming more apparent through investigations of the immune evasion strategies HIV uses to protect conserved regions of the virus. Though, how to overcome these barriers through vaccination remains the greatest challenge.

1.2.1 Obstacles in the Development of an HIV Vaccine

Similar to the reasons that the natural immune response cannot clear HIV as outlined in Section 1.1.4 (Immune Response to HIV Infection), HIV vaccines face the same obstacles. HIV rapidly mutates and recombines under immune selective pressure, leading to extreme antigenic diversity and hypervariability (reviewed in (313, 392)). The virus has adapted ways to overcome both the cellular and humoral immune response, which has enabled HIV to elude vaccine-induced immune control. Furthermore, a number of different clades exist around the world, so an effective vaccine will have to overcome global antigenic diversity. Much of this genetic diversity
arises during replication at the reverse transcription stage, as the error-prone HIV reverse transcriptase enzyme lacks a proofreading mechanism (659, 694). The result is a high mutation rate of the viral genome, and within an HIV-infected individual these single point-mutations in the viral progeny can occur between $10^4$ to $10^5$ times per day (154, 594). This genomic heterogeneity allows viral variants to evade vaccine-elicited immune responses, and requires the development of new epitope-specific CTL and antibody responses. Furthermore, retroviruses have a notoriously high rate of recombination during DNA synthesis (371, 910), which results in the formation of a number of viral quasispecies that add to the diverse population of HIV viral isolates.

Selective pressure by the adaptive immune response and/or antiretroviral therapy causes HIV to mutate in a manner that can evade both the CTL response and antibody-mediated neutralization (86, 636, 687, 869, 885). HIV viral escape is best highlighted in the envelope glycoprotein, which rapidly changes within an individual, between individuals, and between clades at the global level, leading to an unprecedented amount of antigenic diversity (12, 560). Between the various HIV clades, 20-50% nucleotide diversity is observed in the envelope sequence, and within clade M, 25-35% envelope and 15% gag amino acid diversity has been reported between subtypes (407, 814, 826). In addition to genetic diversity, the structure of the protein makes it difficult to target, and sequence hypervariability leads to structural changes during the course of infection (153). The envelope glycoprotein is highly glycosylated, which forms a glycan shield that effectively masks conserved epitopes from the immune response and makes it resistant to neutralizing antibodies by rearranging viral sugar moieties (677, 869). Furthermore, the oligomeric structure of the envelope glycoprotein hides these fundamental immune targets in crypts, rendering them inaccessible to the humoral immune response (892) however, this does not impair viral fitness and the envelope receptor-binding function is maintained (448). These outlined immune escape mechanisms present a scientific challenge in developing an HIV vaccine to target the surface envelope glycoprotein (290), however it does signify an area of vulnerability in the virus and emphasizes the importance of vaccine-induced broadly neutralizing antibodies.

The ability to effectively evaluate novel vaccine concepts is restricted by the current animal models and their associated limitations. Non-human primates (NHPs) have been widely utilized animal models in HIV/SIV vaccine development (293, 579, 596) however, we are
reminded through the results of the Step trial that NHP models are not always positive predictors of efficacy and safety outcomes in humans (reviewed in Section 1.2.3.2 Step Trial) (867). Though closely related to humans, captive NHPs are not exposed to the same pathogens and environmental conditions as humans, which can dramatically impact HIV vaccine safety and efficacy studies. Bone marrow-liver-thymus (BLT) humanized mouse models have recently gained more attention in the HIV field and are being utilized to evaluate a variety of different areas of HIV (94, 215, 230, 434, 568, 804) however, their utility as a model to evaluate HIV vaccine candidates will likely be limited to the early stages of HIV vaccine development and immunogenicity studies.

Limited information is available to define which immune responses would be required to prevent HIV infection or control HIV replication in a vaccine setting. Conventional vaccine approaches have been largely ineffective against HIV’s degree of viral diversity. As such, HIV vaccine researchers are faced with many challenges and have been required to evaluate a number of different vaccine strategies in an effort to limit the HIV pandemic and burden of disease.

1.2.2 HIV/SIV Vaccine Strategies

Much can be learned from the successful vaccines that have been developed to eradicate some of the most pathogenic and harmful viruses. HIV vaccine researchers have put considerable effort into evaluating and applying these licensed vaccine approaches to combat HIV. Classical vaccines have been successful in effectively mimicking the natural immune response associated with viral clearance. With regard to HIV, vaccine researchers are forced to deal with a higher degree of viral complexity, which will require a comprehensive vaccine strategy that targets several aspects of the immune response. The HIV vaccine platforms that are currently under evaluation aim to take advantage of the various properties of the HIV virion (Figure 1-8). The vaccine-induced immune responses, as well as the advantages and disadvantages of these HIV/SIV vaccine strategies will be discussed.
To date, some of the most effective vaccines have utilized whole inactivated and live-attenuated viruses, including smallpox (314), yellow fever (825), hepatitis A (256, 665), poliomyelitis (716, 724), Varicella Zoster Virus (VZV) (813), trivalent Mumps/Measles/Rubella (MMR) (356, 413, 644, 663), and influenza (489). These vaccine approaches have been extensively investigated in an attempt to apply these strategies to HIV vaccine development. In the evaluation of whole inactivated HIV vaccines, inactivated SIV vaccine studies have been shown to be effective at inducing antibody responses, however the antibodies were non-neutralizing and the vaccines were relatively inefficient at inducing CTL responses (389, 468, 569, 570). Live attenuated vaccines (LAVs) remain one of the most effective HIV/SIV vaccine strategies tested in NHP models however, due to inherent safety concerns associated with LAVs they have not been evaluated as HIV vaccines in humans. LAVs induce both a cellular and humoral immune response, and a number of NHP studies, one of which includes a trial conducted by the MacDonald laboratory (878), have highlighted the effectiveness of LAVs in reducing SIV viral load and the progression to Simian AIDS (SAIDs) (16, 193, 390). The immunological mechanism(s) of LAV-mediated protection remains to be elucidated. Although, it
has recently been suggested that the ability of LAVs to establish persistent replication-competent SIV infection in follicular helper CD4+ T cells, located in the secondary lymphoid tissues, may facilitate the maintenance of tissue-based anti-SIV effector memory-differentiated T cell responses that can elicit timely antiviral CTL responses (276). Furthermore, persistent LAVs have also been shown to induce Env-specific ADCC responses that increased over time and were proportional to the degree of LAV replication, suggesting a possible involvement of antibody-mediated protection (17). Unfortunately, the immunogenicity of LAVs is inversely related with the degree of attenuation, thus there is a fine line between generating a safe and effective HIV vaccine for testing in humans (391). The key problem with whole inactivated and live attenuated vaccines is that they have a number of safety concerns, including the risk of reversion to a pathogenic strain of the virus (39, 875). In this regard, the disadvantages live attenuated vaccine strategies clearly outweigh the immunological advantages and as such, until the safety profile of these strategies can be significantly improved these approaches do not present viable vaccine candidates. However, these vaccine approaches are continuing to be studied in an effort to understand the mechanism behind their effectiveness in order to translate these features to other vaccine strategies.

Recombinant protein vaccines have shown to be effective at eliciting humoral immune responses, as observed with the hepatitis B vaccine in which vaccination with a hepatitis B surface antigen elicits protective antibody responses (453, 811). Similarly, vaccination with HIV’s surface envelope protein, gp120, has proven effective at generating vaccine-induced antibody responses (reviewed in Section 1.2.3.1 VaxGen Trials) (296, 315, 638, 640), however the responses have limited neutralization breadth and thus are ineffective at protecting against genetically diverse strains of HIV (262, 296, 639). Due to the difficulty of inducing broadly neutralizing antibodies, much of the focus in the field of HIV vaccine development has been shifted towards evaluating T cell-based vaccines. T cell-based vaccine strategies include vaccination with synthetic peptides encoding T cell epitopes (202, 443) and DNA plasmids expressing vaccine antigens (122). HIV peptides elicit exclusively T cell-based responses, whereas DNA plasmid vaccines can induce both a cellular and humoral immune response that is specific to the inserted vaccine antigens. Of great importance for human clinical trials, these vaccine approaches are safe and pre-existing immunity does not present a problem. DNA vaccines present the initial challenge of getting sufficient DNA into the cells, though several
approaches are being examined such as gene guns, electroporation, etc. Furthermore, to enhance immunogenicity, DNA vaccines are being evaluated in combination with others vaccine approaches, including recombinant proteins or viral vectors, in a prime-boost regimen (146).

Another T cell-based vaccine strategy that has shown much promise is the use of viral vectors to deliver HIV vaccine antigens. The use of recombinant viral vectors has been widely evaluated for veterinary and human application as vaccines for infectious diseases and cancer (225), and was the first HIV vaccine strategy tested in humans (906). In a similar manner to LAVs, recombinant viral vectors are capable of infecting host cells and expressing vaccine antigens to induce both a cellular and humoral immune response, and unlike LAVs, they have a reduced safety concern for pathogenic reversion. A number of HIV/SIV vectors are currently being evaluated at various stages of non-human primate, preclinical, and clinical development (387, 623). Those in clinical development include various serotypes of adenovirus (reviewed in Section 1.2.3.2 Step Trial) (40, 51, 104, 318, 421), poxviruses such as canarypox (reviewed in Section 1.2.3.3 RV144 Trial) (680) and vaccinia (ClinicalTrials.gov: NCT01705223, NCT01783977, NCT01571960) (54, 186, 302, 420, 517, 742), measles virus (ClinicalTrials.gov: NCT01320176) (481, 791), sendai virus (ClinicalTrials.gov: NCT01705990) (418, 513), and vesicular stomatitis virus (ClinicalTrials.gov: NCT01578889, NCT01438606) (176, 388, 697, 698). Viral vector vaccines can be highly attenuated or single-cycle replicating viruses, and thus do not share the same safety risks as live attenuated viruses. However, there can be concerns and complications with the vector itself with respect to safety and immune interference, including pre-existing immune responses to vector antigens that may have unexpected effects (104, 900). An extensive and thorough examination of humoral serostatus, cellular anti-vector immune responses, and the potential for pathogenicity, is imperative prior to evaluating these viral vectors in humans, something that was not previously recognized before the Step trial (reviewed in Section 1.2.3.2 Step Trial).

HIV viral vector vaccines have shown promise in controlling HIV infection, particularly the canarypox RV144 trial (reviewed in Section 1.2.3.3 RV144 Trial), however they have not induced sterilizing immunity. It is clear that a successful HIV vaccine will need a combinatorial approach in which both T cell and B cell mediated systemic and mucosal responses are generated in order to overcome the hypervariability and global genetic diversity of HIV (61, 856). B cells have the potential to induce broadly neutralizing antibodies that block the incoming virus,
whereas antiviral T cells act to control disease by eliminating HIV-infected cells once infection has occurred. To stimulate this combination approach, heterologous prime/boost vaccination regimens are being evaluated and have shown much promise in NHP and human clinical trials (103, 622, 680). This immunization regimen is highly effective at stimulating a robust T cells response by initially priming the memory T cell response, and then subsequently using a boost to further expand the anamnestic T cell response (733). A number of HIV vaccine candidates have employed this strategy with continued and ongoing optimization of the immunization regimen including tweaks to the number of primes/boosts administered, the timing of inoculations, and the combination of recombinant viral vectors, DNA vaccines, and proteins. It is hoped that these efforts will serve to enhance the magnitude of the vaccine-induced immune response and provide protection in human clinical efficacy trials. The HIV vaccine strategies that have advanced to clinical efficacy trials in humans will be described in more detail.

1.2.3 HIV Vaccine Clinical Trials

Although a number of important HIV/SIV vaccine trials have been or are being conducted in humans and NHPs, the focus will be on the four most prominent and relevant human vaccine efficacy trials published to date. The two VaxGen trials were the first Phase III HIV vaccine clinical trials initiated, and evaluated an antibody-mediated recombinant protein vaccine in populations of men who have sex with men (MSM), high-risk women, or injection drug users (Section 1.2.3.1 VaxGen Trials). The Step trial was a T cell-based viral vector vaccine that was evaluated in both men and women in a multicenter study (Section 1.2.3.2 Step Trial). The only human clinical trial to show any protective efficacy, albeit modest, was the RV144 trial; which combined both a cellular and humoral inducing prime-boost strategy in a large population of men and women in Thailand (Section 1.2.3.3 RV144 Trial). These trials each targeted a different arm of the immune response, either alone or in a combinatorial approach, and were mostly conducted over concurrent timeframes with one trial starting before the results of the previous trial was published. Though these vaccines did not generate sterilizing immunity or improve HIV disease progression, they have provided the field of HIV vaccine researchers with much insight into the safety precautions, immunogenicity, and correlates of protection that are the necessary foundation for the development of more effective HIV vaccines in the years to come and thus will be described in detail.
1.2.3.1 VaxGen Trials

The VaxGen trials tested an HIV recombinant protein vaccination strategy designed to elicit broadly neutralizing antibodies that would induce sterilizing immunity. Early non-human primate studies in chimpanzees showed that vaccination with a recombinant HIV-1 glycoprotein gp120 (rgp120) was effective in preventing HIV infection (63, 65, 238). Following intravenous HIV-1 challenge, the rgp120-vaccinated chimpanzees did not show any sign of infection during more than twelve months of follow-up (65), thus efforts were made to translate the non-human primate outcomes into humans clinical studies. VaxGen (AIDSVAX®) is a recombinant HIV-1 gp120 bivalent protein that contains two envelope antigens (62), and has been shown to induce neutralizing antibodies to the incorporated HIV subtypes in vitro (64). Phase I and II human trials demonstrated that the vaccine was both safe and immunogenic in humans (296, 315, 638). The VaxGen Phase III human HIV vaccine efficacy trials were concurrently evaluated in two different populations in an effort to target different envelope subtypes and examine the breadth of the antibody response. For both studies, the primary endpoint for vaccine efficacy was HIV infection, and the secondary endpoints were safety and delayed HIV disease progression.

In 1998, the Vax004 trial examined a bivalent subtype B/B in North America and Europe in a double-blind, randomized, placebo-controlled study that enrolled 5,403 HIV-negative high-risk men and women between the ages of 18-62 years old (262, 340). The participants were given 8 intramuscular inoculations at months 0, 1, 6, 12, 24, 30, and 36. The vaccine was safe, well tolerated, and elicited a robust immune response. All vaccinated participants developed HIV-1 gp120-specific antibody responses, which were able to block gp120 from binding to recombinant soluble CD4, however these responses were unable to neutralize primary HIV strains and the vaccine was not effective at preventing HIV acquisition or modulating HIV disease progression following infection (262, 296). A more detailed analysis of the immune responses generated following AIDSVAX® B/B vaccination demonstrated that peak anti-gp120 antibody responses were inversely correlated with HIV infection and furthermore, vaccinated participants with low anti-rgp120 antibody responses had a higher incidence of HIV infection than the participants in the placebo group (296). Though not definitively proven, the correlation between low antibody responses and HIV incidence was explained as being merely a correlate of susceptibility and not the cause of increased risk of HIV acquisition (296).
Concurrent with the Vax004 trial, a second VaxGen trial was performed in a population of injection drug users (IDUs) in Thailand to examine the rgp120 AIDSvax® vaccine in a population with diverse genetic subtypes of HIV, in which the most common circulating HIV-1 envelope subtypes are B and E (403, 472, 803). In 1999, the Vax003 trial evaluated a rgp120 bivalent subtype B/E in a double-blind, randomized placebo-controlled trial and enrolled 2,546 HIV-negative male and female IDUs between the ages of 20 and 60 years old (639). The vaccine was administered intramuscularly at months 0, 1, 6, 12, 18, 24, and 36. The results of the study showed that the vaccine was not effective and did not alter HIV disease progression following HIV infection. With respect to the study endpoints, no significant differences were observed between the vaccinated and placebo participants, though the vaccine was deemed safe with no adverse effects (639).

The VaxGen trials were completed successfully without any safety concerns, however neither vaccine induced neutralizing antibodies against primary HIV strains, nor did they display any significant efficacy for preventing HIV infection or improving HIV disease progression. Of the two vaccines, the AIDSvax® B/B had a higher vaccine efficacy (6%) than the AIDSvax® B/E (0.1%), though neither of which were statistically significant compared to the placebo groups (262, 639). The antibodies generated in response to the vaccine were only able to neutralize laboratory-adapted strains of HIV-1 and were unable to neutralize primary isolates of HIV-1, as observed in both the human and non-human primate trials (65, 296), which confirms previous reports (508, 561) and makes the vaccine ineffective in a real-world setting. These first Phase III efficacy trials did provide valuable insight with respect to epidemiology, high-risk populations, as well as compliance, and behavioural influences that helped to direct subsequent and ongoing clinical trials.

1.2.3.2 Step Trial

In contrast to the antibody-focused VaxGen trials, the Step trial explored a T cell-based vaccine strategy that utilized adenovirus serotype 5 (Ad5) as a recombinant viral vector to deliver the vaccine antigens. A number of non-human primate vaccine studies, conducted in rhesus macaques, demonstrated that a DNA prime and Ad5 boost with a replication-incompetent Ad5 vector containing HIV or SIV proteins, elicited T cell-mediated immune responses to the vaccine antigens, and following SIV challenge, resulted in reduced peak viral load and in some
cases lowered set-point viral loads (120, 465, 752, 882). The promising immunogenicity and efficacy observed in the NHP trials led the way to a Phase I human clinical trial, in which the Ad5 vector expressing HIV proteins was deemed safe and immunogenic in humans (121, 661). The subsequent Phase II trial began in 2004, termed the Step trial, in a collaborative effort between Merck and Co., Inc., National Institute of Allergy and Infectious Disease (NIAID) at the National Institute of Health (NIH), and the HIV Vaccine Trials Network (HVTN). The double-blinded randomized study comprised 3,000 HIV-negative participants in a multicenter study with enrolment in the Caribbean, North America, Australia, and South America (104). The men and women enrolled in the study were between 18-45 years of age and were selected to be at high risk for acquiring HIV based on their recorded risk behaviours (104). The vaccine was a recombinant replication-defective adenovirus serotype 5 (Ad5) vector with HIV clade B Gag, Pol, and Nef proteins incorporated, termed MRKAd5. The vaccine (MRKAd5) or placebo (vaccine diluent with no Ad5 vector) was given in three intramuscular inoculations at weeks 0, 4, and 26, and the participants were monitored for HIV infection with testing performed every 6 months (approximately) (104). The primary endpoints of this study were to demonstrate safety, tolerability, and efficacy of the vaccine, with efficacy defined by prevention of HIV-1 infection, and reduced viral load following infection (104).

In 75% of the vaccinated participants, the vaccine elicited both CD8+ and CD4+ T cell responses to the HIV vaccine antigens (104, 523). However, the vaccine-induced cell-mediated immune response was not effective at preventing HIV acquisition, and unlike the NHP trials, it did not lower peak viral load following infection (104). The trial was prematurely halted at the first interim review when the Data Safety Monitoring Board (DSMB) identified increased HIV acquisition rates in the vaccine group compared to the placebo group. A subset of vaccinated participants who had pre-existing Ad5 antibody titers and who were uncircumcised appeared to have a higher risk of acquiring HIV during the first 18 months following vaccination (104, 231). Running concurrent to the Step trial was another Phase II study, the Phambili trial (HVTN 503), which was testing the same MRKAd5 vaccine in a South African-based study consisting of 801 participants. The Phambili trial was halted prematurely during the enrollment and vaccination stage due to the findings in the Step trial (318).

There has been much debate regarding the causative mechanism for the increased HIV acquisition in this subset of the vaccine arm. In an effort to elucidate the mechanism behind the
observed increase HIV acquisition in the vaccinees, a post-hoc analysis examined a number of variables in a comparison between the individuals who acquired HIV (cases) and those that did not (non-cases) (523). Because all but one of the participants that acquired HIV were men, more specifically men that have sex with men (MSM), the post-hoc analyses were focused on the MSM population. Vaccine-induced HIV-specific T cell responses were not modulated in the cases, as determined by response rate, magnitude, and polyfunctionality; furthermore the number of circulating CD4\(^+\) T cells expressing CCR5 (HIV co-receptor) were comparable between the cases and non-cases (523). With respect to vector-based immunity, the cases had lower circulating Ad5-specific CD4\(^+\) T cell responses compared to the non-cases. It is hypothesized that these CD4\(^+\) T cells migrated to the mucosal sites resulting in an increase in HIV target cells at the site of infection, which in uncircumcised men provided a larger surface area for the target cells to home (60, 523). Though tissue samples were not available to examine activated CD4\(^+\) T cells at the mucosal sites, a subsequent *in vitro* analysis using blood samples collected from 20 healthy donors was performed to test the above hypothesis. The results demonstrate that individuals with pre-existing Ad5 immunity had a positive correlation between Ad5 antibody titer and the number of CD4\(^+\) T cells expressing the gut homing molecules α4β7 and CCR9, suggesting that vector-induced memory CD4\(^+\) T cells migrated to the site of HIV infection leading to increased susceptibility (60). Furthermore, it was also shown that expanded Ad5-specific CD4\(^+\) T cells expressed increased levels of CCR5 and were more susceptible to HIV infection upon re-stimulation with Ad5 (60). The prevailing evidence suggests that activated Ad5-specific T cells at the mucosal sites impacted the results of the Step trial however, because mucosal samples were not available, this theory cannot be confirmed nor denied. Since the results of the Step trial were published, a number of other studies have refuted this theory (134, 185, 375, 595, 792), and thus up to the present time no causal link exists.

The NIAID and HVNT have very recently announced in a press release that another adenovirus-based vaccine trial, HTVN 505, has been discontinued and the vaccine regimen has been deemed ineffective (587). Akin to the Step trial, the recently discontinued HTVN 505 was a Phase IIb trial examining a prime/boost vaccination regimen of 3 primes with a recombinant DNA plasmid expressing HIV Gag, Pol, Nef, and Env, followed by a single boost with an attenuated adenovirus type 5 (Ad5) vector expressing HIV Gag, Pol, and Env (ClinicalTrials.gov: NCT00865566). In an effort to evaluate a similar vaccination strategy to the
Step trial in a population that is not at risk for vaccine-induced increased HIV acquisition, the trial began in 2009 and was conducted in the United States in a population of 2,504 HIV-negative MSM and transgender women who have sex with men that were both Ad5 seronegative and circumcised. The preliminary results show that the vaccine did not prevent HIV infection or have an impact on reducing viral load following infection; and although not statistically significant, more cases of HIV infection were observed in the participants that received the vaccine compared to those that received the placebo (587). The investigators are currently working to elucidate the reasons for which this vaccine regimen was unsuccessful and will continue to follow the participants in hopes of gaining additional information on the trial outcomes (587). The take-home message from the adenovirus-based HIV vaccine trials is that adenoviruses, particularly serotype 5, do not represent viable and safe vectors for an HIV vaccine, and NHP models are not always positive predictors of vaccine safety. In moving forward, researchers need to consider the impact of seropositivity, pre-existing immunity, and vector-based immunogenicity when developing and evaluating alternative viral vector candidates for an HIV vaccine.

### 1.2.3.3 RV144 Trial

Despite disappointment following the failures of the Step trial, the field was reinvigorated by the optimistic results of the RV144 trial that began in 2003 and was published in 2009. The RV144 trial was a large Phase III double-blinded randomized study comprised 16,395 HIV-negative participants in a multicenter study with enrolment in Thailand (680). The study was a collaborative effort between Thailand (Thailand Ministry of Public Health) where the trial was conducted, and the United States [US Army, and National Institute of Allergy and Infectious Disease (NIAID)] who sponsored the trial. Men and women between 18-30 years of age whom were at community risk for acquiring HIV were enrolled in the study. In an effort to induce both a cellular and humoral immune response, the vaccine was a prime-boost vaccination with a recombinant canarypox vector expressing HIV Gag, Pol, and Env proteins (ALVAC-HIV) delivered as a prime and the VaxGen recombinant gp120 subtype B and E bivalent protein (AIDSVAX B/E®) was given as a boost. The vaccination regimen took place over 6 months and included four priming inoculations with ALVAC-HIV at weeks 0, 4, 12 and 24, and two booster inoculations with AIDSVAX B/E® at weeks 12 and 24. Participants were monitored for HIV
infection every 6 months for 3 years. The primary endpoints for the study included prevention of HIV-1 infection, and upon infection, the impact of the vaccine on HIV-1 viral load (680).

Prior to the RV144 trial, both the prime and boost components were evaluated separately and in combination, and were deemed safe and immunogenic in a number of Phase I/II trials conducted in Thailand and the United States (58, 59, 411, 590, 591, 638-640, 710, 820, 827). In the Phase III RV144 trial, the vaccination induced modest cellular immunogenicity with 19.7% of the vaccinees displaying HIV-specific T cell responses to the prime antigens (680), which was considerably lower than the 75% of vaccinees responding in the Step trial (104, 523). With regard to the humoral immune response, the vaccine group exhibited robust HIV-specific antibody and lymphoproliferative responses to the boost protein (680). Most interestingly, the major comparison between the failed Step trial and the RV144 trial, was the safety, tolerability, and efficacy of the RV144 vaccine (641, 680). Overall, the vaccine regimen was 31.2% effective at reducing HIV infection in the vaccinated participants but had no impact on HIV disease progression in the individuals that did become infected, as there was no significant difference in HIV viral load or CD4+ T cell counts in the vaccines compared to the placebo group (679, 680). Although the vaccine did not reduce HIV viral load in the periphery, recent studies have shown that the vaccinated individuals that became infected may have reduced viral load in the genital mucosa (679), which may have an added benefit of reducing HIV transmission. Following the optimistic protective efficacy results, a number of post-hoc studies were performed to further examine the observed protective efficacy and elucidate the immune correlates of protection. The most intriguing finding was that vaccine-induced IgG antibodies that bind to the variable 1 and 2 (V1V2) region of the HIV-1 envelope glycoprotein were shown to positively correlate with protection (343, 410, 466, 626, 920). Whereas Env-specific plasma IgA antibodies were shown to negatively correlate with protection but not increase the risk of infection (343). Although not shown to be correlates of protection, higher levels of antibody-dependent cell-mediated cytotoxicity (ADCC) and Env-specific CD4+ T cell responses were observed in the vaccine group (80, 343, 680). Moreover, the vaccine-induced CD4+ T cell responses were preferentially directed towards HIV envelope V2 epitopes (203), suggesting that these immune responses may be playing a protective role.

The partial protection observed in the RV144 trial represents a major step forward and the most promising HIV vaccine trial to date. Though the efficacy was not substantial enough to
warrant licensing of the vaccine, the results of this trial will serve to inform future HIV vaccine trials. The vaccine strategy is being reevaluated in non-human primates in an attempt to recapitulate the results of the RV144 trial and gain more insight into the correlates of protection. Preliminary data from one study that mimicked the RV144 vaccination strategy in rhesus macaques revealed that the vaccination regimen upregulated innate immune responses and induced NK cell activation (477). Furthermore, these activated NK cells had an increased frequency at mucosal sites, with a particular NK22 phenotype that expressed a gut homing marker, suggesting a role in mucosal immunity (477). In addition to these important non-human primates studies that may help to define correlates of protection, a number of follow-up human clinical trials are ongoing with modifications to the vaccination regimen, sample size, tissue sampling, etc. In the Thai trial, the efficacy of the RV144 vaccine peaked early at 6 months following the final vaccination (60.5% efficacy), however the efficacy waned rapidly resulting in an overall protective efficacy of 31.2% at the end of the study (3 years following the final vaccination) (680, 692), which suggests that ongoing boosting may help to increase and maintain the protective efficacy of the vaccine. It is hoped that researchers will be able to improve the modest efficacy of this vaccine in order to achieve either sterilizing immunity or a functional cure. Either way, these studies will continue to provide valuable information for the vaccine field on immune correlates of protection that will help drive HIV vaccine development.

1.2.3.4 Ongoing HIV Vaccine Clinical Trials

While the RV144 trial renewed hope within the field of HIV vaccine development, the effectiveness was modest at best and there is still no effective HIV vaccine. According to the IAVI Clinical Trial Databases (377), there are currently 41 ongoing or scheduled HIV vaccine clinical trials worldwide. Thirty-two of these trials are testing viral vector vaccines, almost all of which are adenovirus- or poxvirus-based vectors. With the failed Ad5 vector trials, much attention has been placed on the RV144 follow-up trial (RV305). RV305 is beginning enrollment to test a modified vaccination schedule in hopes of enhancing the vaccine-boosting arm to increase efficacy (ClinicalTrials.gov: NCT01435135). The limitations to both adenovirus and poxvirus vectors are that they are attenuated or non-replicating so the antigen provided by these vectors wanes over time. Thus, these vaccines rely on a central memory-derived anamnestic response to elicit effector memory T cells, which is too late for HIV, as once the virus moves from the local site of infection to the periphery there is no chance of clearing the virus with the
adaptive immune response. Thus, instead of making incremental tweaks to the current vaccine approaches, other vaccine strategies need to be evaluated. Though central memory T cells have great proliferative potential, they reside predominantly in lymph nodes, thus it takes longer for them to respond to antigen and act in a killing capacity. In contrast, effector memory T cells can be predominantly found in the mucosa, and offer an almost immediate first responder phenotype at the portal of HIV entry (vaginal, rectal mucosa). Hence, a successful HIV vaccine will likely require long lasting memory with the capacity to respond quickly at the level of the mucosa, and there is a growing view that chronic viral vectors may be superior in their ability to induce and maintain an effector memory T cell response. We hypothesize that in order to overcome the delayed anamnestic effector response, a successful HIV vaccine must induce persistent antigen expression to maintain primed effector memory cells, which may act to remove the virus immediately at the site of HIV entry and before it can disseminate to the periphery. One way to overcome this limitation and generate persistent antigen exposure is through the use of herpesviruses as viral vectors, which is the longstanding focus of the MacDonald laboratory. We anticipate that using chronic reactivating viruses, such as herpesviruses, as HIV vaccine vectors will provide a durable and persistent in situ mucosal peripheral T cell effector memory response by eliciting protective effector cells at the portal of HIV entry and in a timeframe most appropriate to inducing clinically relevant protection.

1.3 Herpesviruses as Viral Vectors for an HIV/SIV Vaccine

1.3.1 Rationale for Herpesvirus-Based Viral Vectors

Herpesviruses establish life-long infection in the host, with evidence of repeated subclinical reactivation and immunogenicity, which could theoretically facilitate episodic expression of vaccine antigens. This ability of the virus to “self-boost” represents a significant benefit over other HIV vaccine vectors currently in clinical trials and may help address the key issue of transient and waning vaccine immunity (previously reviewed in Section 1.2.3 HIV Vaccine Clinical Trials). The rationale for examining herpesviruses as HIV viral vectors stems from two areas that have been investigated by the MacDonald laboratory, which include the natural immunity observed in highly exposed, persistently seronegative Kenyan female sex workers (358, 414, 416), and the efficacy of SIV live attenuated vaccines (LAVs) in non-human primates (878). The observation, in these two highly distinct research areas, that persistent
antigen exposure at mucosal sites might be protective from HIV/SIV acquisition sheds light on potential immune correlates of protection for an effective HIV vaccine.

Low levels of HIV-1 virus have been found in the mononuclear cells of highly exposed persistently seronegative (HEPS) individuals (915), and despite remaining HIV seronegative, female sex workers exhibit HIV-specific cellular and humoral immune responses in both the blood and genital tract (267, 415, 416, 703). Furthermore, it has been shown that pre-acquisition HIV-specific CD8+ T cell responses in HEPS female sex workers that became infected with HIV were associated with a reduced HIV viral load set-point (414). Thus, we suggest that persistent exposure may generate and prime HIV-specific T cell responses that might serve to reduce the burden of HIV disease following infection. Similarly, our group has recently observed a similar phenomenon in persistently seronegative cynomolgus macaques that were repeatedly challenged intra-rectally with low-dose SIV and did not seroconvert (Haq, K. et al, unpublished). Thus, priming the host immune response by repeated exposure may have protective effects on acquisition. Likewise, live attenuated vaccines (LAVs) have shown that persistent vaccine replication drives protective efficacy and the degree of attenuation corresponding to viral replication impacts the effectiveness of LAVs (previously discussed in Section 1.2.2 HIV/SIV Vaccine Strategies) (391). This persistent replication observed in LAVs has been shown to induce vaccine-specific effector-differentiated memory T cells at the site of vaccine replication (276, 498). Furthermore, the protective efficacy of LAVs is correlated with the level of persistent LAV replication, as highly attenuated vaccines with limited or single-cycle replicative capacity were not as effective at reducing SIV viral load (171, 212, 244, 276, 385, 891). Thus, these research areas provide support for low-level ongoing viral exposure or replication driving persistent antigen stimulation, which may trigger the immune response to maintain a primed and activated effector memory population in the mucosa. Moreover, this may be facilitating long-lived protection and therefore offers an avenue of exploration for employing a persistent viral vector approach to HIV vaccine design.

In this regard, a closer look at the importance of a persistent HIV/SIV-specific effector memory response at the mucosa is warranted. In both HIV and SIV mucosal infection, there is a small “window of opportunity” in which the immune system can act to control or limit viral replication and contain the virus before tissue dissemination. This window of opportunity occurs immediately following mucosal infection, as once the virus crosses the mucosal barrier it beings
local propagation and expansion in tissue resident CD4$^+$ T cells (Figure 1-9) (329, 369, 393, 549, 555, 912). During the earliest stages of infection, the virus is vulnerable to immune control, as only a small number of CD4$^+$ T cells are infected with a founder population of genetically homogeneous viral isolates, thus limiting the capacity for viral escape (216, 422, 886, 918). However, less than 7 days post-infection, sufficient viral expansion allows the virus to disseminate to the periphery where there is a burst in viremia, initiation of massive CD4$^+$ T cell depletion, and the establishment of a latent viral reservoir, making it virtually impossible for the immune system to clear the virus (329, 393, 549, 555).

**Figure 1-9 Mucosal HIV/SIV Infection** [Republished with permission of Annual Reviews, Inc., adapted from (555); permission conveyed through Copyright Clearance Center, Inc.]

With this in mind, HIV/SIV vaccinologists are required to take a closer look at effector T cell responses with respect to the timing and location of the response, and the subsequent impact on viral load (Figure 1-10). It is apparent that in unvaccinated macaques (gray line), *de novo* effector memory T cell responses occur between 21 and 28 days post-infection, once the virus has disseminated to the periphery and systemic infection has occurred (684). This delayed initiation of the natural immune response occurs subsequent to the burst in viremia and peak viral load, which is effectively too late (634, 684). Similarly, conventional HIV/SIV vaccines (blue
line), which include replication-incompetent viral vectors, do not generate persistent antigen stimulation and thus upon HIV/SIV infection, the immune response must rely on an anamnestic central memory response for effector memory T cell differentiation. This response typically occurs between 14 and 21 days post-infection and like the unvaccinated response, effector memory T cells appear after systemic infection is established and when the virus has already amplified to reached peak viral load (Figure 1-10) (634). Though conventional HIV/SIV vaccines have been effective at reducing SIV viral load in non-human primates studies, they have been largely ineffective in human clinical trials, as reviewed in Section 1.2.3 (HIV Vaccine Clinical Trials). Alternatively, herpesvirus infections induce a broad cellular and humoral immune response both at the mucosa and systemically. Furthermore, these viruses establish lifelong latency with periodic reactivation enabling persistent immune exposure of vaccine antigens, which may facilitate a primed and differentiated effector memory response at the portal of HIV entry (vaginal or rectal mucosa). It is proposed by our group and others, that the periodic reactivation afforded by a herpesvirus-based viral vector vaccine (purple line) will generate persistent antigen stimulation, which will drive the primed anti-HIV/SIV effector memory population in the mucosa. Thus, upon mucosal HIV/SIV infection, the vaccine-induced effector memory T cells will be able to act during the “window of opportunity” to control or limit viral replication at the site of infection and most importantly prior to peripheral dissemination (Figure 1-10) (634).

It is critical to acknowledge that this idea represents a T cell-based vaccine strategy with the goal of lowering HIV/SIV viral load and reducing pathogenicity, thus sterilizing immunity is not expected. As will be discussed in greater detail in the Discussion (Section 5.2.2 The Development of CyCMV as an HIV/SIV Vaccine Viral Vector), this is only one arm of a proposed HIV vaccine strategy and an additional component, such as a protein boost, will be required to induce a humoral immune response that may neutralize the virus and thus work in concert with a herpesvirus-based viral vector in hopes of eliciting sterilizing immunity.
1.3.2 Herpesviruses as Viral Vectors

Mammalian herpesviruses, from the *Herpesviridae* family, are ancient DNA viruses that have evolved over the last 200 million years (524). Herpesviruses are large linear double-stranded DNA viruses that establish lifelong infection in their host and human herpesviruses are categorized into three subfamilies, alpha, beta and gamma (627). The alphaherpesvirus family is comprised of Herpes Simplex Virus-1 (HSV-1 or HHV-1), Herpes Simplex Virus-2 (HSV-2 or HHV-2), and Varicella Zoster Virus (VZV or HHV-3). The betaherpesvirus family includes cytomegalovirus (CMV or HHV-5) and roseoloviruses (HHV-6A, HHV-6B, and HHV-7), while the gammaherpesviruses include Epstein-Barr Virus (EBV or HHV-4) and Kaposi’s Sarcoma-associated herpesvirus (KSHV or HHV-8). The alphaherpesviruses infect mucosal epithelial cells and are known to latently infect the neurons, causing oral and genital herpes (HSV-1, HSV-2), or chickenpox and herpes zoster (VZV) (159, 695). Betaherpesviruses have a broad cellular and tissue tropism, and primarily infect cells of myeloid origin and then latently infect lymphocytes,
particularly T cells (553, 896). Although typically asymptomatic, CMV can cause CMV mononucleosis or more severe diseases in immunocompromised host (reviewed in Section 1.3.3 Cytomegalovirus), whereas HHV-6/7 causes a skin rash known as roseola (553, 896). The gammaherpesvirus family comprises lymphotrophic viruses that replicate in epithelial (EBV) or endothelial (KSHV) cells and latently infect B cells and epithelial cells, resulting in infectious mononucleosis (EBV), or angioproliferative and inflammatory lesions known as Kaposi’s sarcoma (KSHV) (285, 688). In immunocompromised hosts, gammaherpesviruses have a high oncogenic potential, which can lead to the development of lymphomas and carcinomas, such as Burkitt’s lymphoma, nasopharyngeal carcinoma, CNS lymphoma with EBV infection (688), and Kaposi’s sarcoma with KSHV infection (285).

Given the high seroprevalence and lifelong infection induced by herpesviruses, it may be advantageous for a herpesvirus-based viral vector to be able to superinfect seropositive individuals. The majority of the herpesviruses, including HSV-1, HSV-2, VZV, CMV, EBV, and KSHV have been shown to have the capacity to superinfect (4, 69, 127, 543, 593, 678, 722, 859). As a result, a number of herpesviruses are being evaluated as SIV viral vectors in NHP models, including human HSV-1 (417, 476, 571), human and simian VZV (788, 877), rhesus macaque CMV (discussed in greater detail in Section 5.1.2 Rhesus Macaque Cytomegalovirus-Based Viral Vector) (334, 338), and a close relative of KSHV, rhesus monkey rhadinovirus (RRV) (70). HSV-1 is currently being evaluated as a vaccine vector and has shown promise in non-human primate studies (417, 571), however the clinical symptoms associated with HSV-1 infection merit caution. In this regard, safety mutations are being incorporated into the HSV-1 vector in an attempt to improve the safety profile of this candidate vaccine (476). Alternatively, our laboratory has invested considerable effort into developing and evaluating VZV and CMV as viral vectors for an HIV/SIV vaccine.

Varicella Zoster virus and cytomegalovirus are non-oncogenic, non-integrating viruses with rare and limited clinical symptoms in immunocompetent individuals, thus presenting the highest safety profile within the herpesvirus family (159, 553, 627). VZV and CMV share a number of features that make them viable HIV viral vectors, and yet they differ in their own respect, as described in Table 1-1. Both viruses are large DNA viruses, which makes them amenable to manipulation and the insertion of exogenous vaccine antigens when cloned as a Bacterial Artificial Chromosome (reviewed in Chapter 4). Furthermore, they encode a number of
functional and non-functional genes that can be removed to attenuate the virus for safety reasons and/or to facilitate the insertion of vaccine antigens (159, 553). It has been documented that human CMV encodes 68 ORFs that are dispensable for viral replication in vitro (233), which could be removed during the insertion of antigens without impeding viral growth kinetics, though this would need to be subsequently assessed in vivo as the function of the majority of these genes remains to be elucidated. With respect to pathogenesis, VZV leads to a tissue-based infection with a narrow neurotropism, whereas CMV infection targets a broad cellular and tissue tropism (Table 1-1). In comparison to VZV, CMV is a circulating virus that induces a strong T cell response with the presence of both peripheral and tissue-based effector memory cells (discussed in Section 1.3.3.3 Immune Response to Cytomegalovirus Infection). CMV-specific effector T cells continue to clonally expand with host age, in contrast to VZV in which VZV-specific memory T cell responses decrease with age, a phenomenon thought to be associated with the location of latent viral infection (35, 427, 592). The high frequency of reactivation and viral shedding associated with CMV infection may enable a high rate of boosting of vaccine antigens and the induction of an effector-differentiated T cell response, or alternatively it may place increased demands on the host immune system which may lead to immunosenescence (reviewed in Section 5.1.3 Limitations and Safety Concerns to CMV-Based Vectors). In this regard, the advantages and potential limitations of CMV-based HIV/SIV viral vectors will need to be evaluated in a NHP model prior to human clinical trials.

Table 1-1: Comparison of VZV and CMV as HIV/SIV Vaccine Vectors

<table>
<thead>
<tr>
<th>Features</th>
<th>Varicella Zoster Virus (VZV)</th>
<th>Cytomegalovirus (CMV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification</td>
<td>Alphaherpesvirus</td>
<td>Betaherpesvirus</td>
</tr>
<tr>
<td>Size</td>
<td>Smallest herpesvirus</td>
<td>Largest herpesvirus</td>
</tr>
<tr>
<td>Tropism</td>
<td>Neurotropic</td>
<td>Broad cellular tropism</td>
</tr>
<tr>
<td>Latency</td>
<td>CNS</td>
<td>Lymphocytes and myeloid cells</td>
</tr>
<tr>
<td>Reactivation</td>
<td>Less frequent</td>
<td>Highly frequent</td>
</tr>
<tr>
<td>Effector memory</td>
<td>Tissue-based EM cells</td>
<td>Circulating and tissue-based EM cells</td>
</tr>
<tr>
<td>Features</td>
<td>Varicella Zoster Virus (VZV)</td>
<td>Cytomegalovirus (CMV)</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Viral shedding rate</td>
<td>Low</td>
<td>Intermittent</td>
</tr>
<tr>
<td>Circulating virus</td>
<td>Absent</td>
<td>Circulating</td>
</tr>
</tbody>
</table>

One of the benefits for using VZV as an HIV viral vector is that VZV is already an FDA licensed vaccine (Varivax™ and Zostavax™ by Merck & Co., Inc.; Varilrix™ by GlaxoSmithKline Inc.), and thus has already been deemed safe for immunization in both children and adults (31, 34, 138, 454, 541, 608, 757, 841, 870). The pre-existing global licensure of VZV breaks down many barriers with moving forward in the development of VZV as an HIV vaccine viral vector. Alternatively, there are a number of safety concerns associated with CMV-based HIV vaccines, which may result in extensive clinical safety trials in both NHPs and humans prior to proceeding to a HIV protective efficacy clinical trial (reviewed in Section 5.1.3 Limitations and Safety Concerns to CMV-Based Vectors). With the pros and cons of each vector in mind, we are comparing and contrasting these two viruses in a side-by-side/concurrent evaluation of VZV and CMV as viral vectors, in hopes of taking advantage of the tropism differences for a dual vaccine that could be complimentary in an HIV vaccine setting. The primary aim of this thesis is the development of CMV as a viral vector for an HIV/SIV vaccine and thus will be the primary focus from this point forward.

1.3.3 Cytomegalovirus

Cytomegalovirus (CMV) is a ubiquitous host-restricted pathogen that is horizontally transmitted through genital secretions, breast milk, saliva, urine, or through transplantation, and is vertically transmitted via placental transfer (283, 630). CMV has a broad tissue tropism and primarily replicates in mucosal sites, including salivary glands, respiratory tract, gastrointestinal tract, and genital tract (451, 553). Human cytomegalovirus (HCMV) infection is typically asymptomatic in healthy individuals, however in some individuals HCMV can cause clinical symptoms including EBV-like mononucleosis (283, 766). Conversely, when HCMV replication cannot be controlled by the cellular immune response it can cause severe disease, such as in neonates when the immune system is immature, or in immunocompromised individuals for example transplant patients, chemotherapy patients, or AIDS patients (241, 451, 706, 709). In
pregnant women, primary infection, recurrent infection, or superinfection can lead to congenital HCMV infection in the fetus, the leading cause of hearing loss, vision loss, and infectious brain damage in the United States (81, 83, 164, 480, 785, 874). Congenital HCMV infection affects the central nervous system and the major sensory organs of the infant, leading to a number of possible symptoms namely, microcephaly, reduced muscle strength, lethargy, hearing impairment, and seizures (82). In immunosuppressed patients, common clinical effects include, but are not limited to, retinitis, pneumonitis, hepatitis, esophagitis, gastritis, and enterocolitis (283, 766).

To limit pathogenesis in immunocompromised patients, prophylactic treatment with antiviral drugs is routinely administered prior to immunosuppression in the case of transplant or chemotherapy patients and is effective at reducing HCMV infection and disease (10, 46, 482, 544, 883). Furthermore, efforts have been made to generate a preventative HCMV vaccine using attenuated HCMV strains (239, 581). The most well documented attempt was a live attenuated HCMV vaccine, based on the Towne 125 strain, which was not successful at preventing HCMV infection, as the degree of attenuation was too severe to mount a sterilizing immune response, however the vaccine was effective at preventing or reducing disease in immunosuppressed individuals (8, 300, 643, 647-650). Similar strategies to those being examined for an HIV vaccine, as outlined in Section 1.2.2 (HIV/SIV Vaccine Strategies), are being evaluated for CMV vaccines (305, 306, 672). As previously discussed in relation to the pre-existing licensed VZV vaccine, a licensed safe and effective CMV vaccine would be advantageous in developing a CMV-based HIV vaccine and could also be used for co-CMV/HIV immunization.

Globally, HCMV seroprevalence ranges from as low as 45% to as high as 100%, with the highest being in the developing world (114). A number of studies have shown a sex bias in CMV seroprevalence with marginally higher prevalence in woman than in men (55, 347, 789). Furthermore, with respect to age, the prevalence of CMV increases with age, as shown in an epidemiology study conducted in the United States that examined CMV serostatus in the ≥ 80 years of age cohort and found that 90.8% of the individuals were CMV seropositive (789). Limited data is available on National HCMV seroprevalence in Canada, however a small number of provincial studies have been conducted in either whole or specified populations. The province-wide HCMV seroprevalence in Nova Scotia was reported to be greater than 22% (240), and in other provinces it was observed that the percent seropositive in the specific population
groups examined was much higher with 57% in a cohort of female day care educators in Montreal (398), 67% in day care providers from childcare centers in Toronto (264), and 78% in pregnant women in British Colombia (850). To estimate the Canadian National HCMV prevalence, the United States may be used as a best comparison, which has a National HCMV prevalence of approximately 50% or greater, depending on the region (55, 553, 789). It should be noted that the epidemiological data does not represent the population on a whole as it is taken from a biased population of blood donors, study participants, employees, etc., and thus is only an approximation and likely represents an underestimate of seroprevalence. Thus, with a high HCMV seroprevalence rate, it is imperative that an HCMV-based HIV vaccine is able to superinfect seropositive individuals (also reviewed in Section 5.1.3 Limitations and Safety Concerns to CMV-Based Vectors). Fortunately, though HCMV elicits a strong cellular immune response, it is not sufficient to prevent re-infection with a different strain of HCMV, as superinfection with multiple strains of HCMV has been well documented in both immunocompetent (127, 543, 593) and immunocompromised adults (151, 226, 781, 818), as well as in children (45).

1.3.3.1 Cytomegalovirus Structure and Genome

Human cytomegalovirus is a Class E structure with 4 genome isoforms and an approximate diameter of 200 nm (295, 553). HCMV is an enveloped virus comprised of at least 10 surface glycoproteins, six of which have been characterized and include gB, gH, gL, gM, gN, gO (99, 553, 635, 670) (reviewed in Section 3.5.7 Surface glycoproteins). Below the envelope phospholipid membrane is the proteinaceous tegument layer that is filled with proteins that surround and coat the nucleocapsid, an icosahedral structure (approximately 110 nm) that contains the linear dsDNA viral genome (Figure 1-11) (133, 294). In addition to their structural role, tegument proteins are highly immunogenic and have a number of different functions in the various stages of viral replication including the early inhibition of host cell cycling (99, 451).
Human cytomegalovirus is 230 kbp in length and encodes over 200 ORFs, making it the largest and most complex herpesvirus (553). Various lab-adapted and clinical strains of HCMV have been isolated and sequenced, most notably are AD169 (131), Toledo (572), Towne (233), and Merlin (218). Furthermore, there are a number of clinical strains that have been cloned as bacterial artificial chromosomes (BACs), such as TB40/E (763) and TR, PH, FIX (VR1814) (572). The HCMV genome consists of terminal repeats, terminal repeat long (TRL) and terminal repeat short (TRS), at either end of the genome and is organized by unique long (UL) and unique short (US) segments separated by an internal repeat (IR) region (Figure 1-12) (99). The UL region encodes the most conserved genes, which includes the 40 core genes, that are conserved amongst all alpha-, beta-, and gamma-herpesviruses (553). The betaherpesvirus family, consisting of CMV, HHV-6 and HHV-7, diverged as a separate herpesvirus family over 170 million year ago, with the subsequent divergence of CMV 60 million years thereafter (524, 525). Cytomegaloviruses have co-evolved with their host species resulting in highly species-specific viruses (524), and the continued virus-host co-evolution has resulted in the acquisition of a series of host-specific genes that are of benefit to the virus. A number of these genes lead to enhanced viral replication and during primary infection, the incorporation of anti-apoptotic and immunoevasin genes inhibit the host immune response from clearing the virus (3). Specific CMV genes and their functions will be described in great detail in Chapter 3, more specifically Table 3-1 and Table 3-2, and thus will not be described in this section.
In addition to having broad tissue tropism, HCMV has an extensive target cell tropism and infects a number of different cell types, including cells of myeloid origin, fibroblast cells, epithelial cells, endothelial cells, neuronal cells, smooth muscle cells, and hepatocytes (553, 764, 765). During lytic replication, HCMV infects cells of myeloid origin and HCMV establishes latency in these cell types as well as in lymphocytes (383, 735). Furthermore, differentiation of latently infected myeloid cells into their mature phenotypes leads to HCMV reactivation (383, 774), and both endothelial cells and lymphocytes are involved in the spread of HCMV throughout the various tissues (451, 735, 765). In vitro, CMV replicates in fibroblast and endothelial cells, and like all betaherpesviruses, CMV has characteristically slow growth kinetics in culture (553). CMV infected cells exhibit cytopathic effect (CPE) characterized by the presence of enlarged rounded cells in the form of clusters (451). In comparison to primary clinical isolates, lab-adapted viruses routinely eliminate genes that are not required for propagation in vitro, particularly tropism genes, which do not affect viral growth kinetics in culture (124, 233) (reviewed in Section 3.5.5 Tropism Genes).

The HCMV replication cycle takes 48-72 hours to complete, the stages of replication and the associated genes are illustrated in Figure 1-13 (553). The surface glycoproteins are essential for viral attachment, fusion, and penetration of the target cell, with an additional role in cell-to-cell spread (100). Glycoprotein B binds to heparin sulfate proteoglycans on the surface of the target cell and this tethering allows the viral entry process to begin (167, 451). In the second stage of entry, gH, gL, and gO fuse the viral envelope with the host cell membrane, in a pH-independent manner, enabling the virus to penetrate the cell membrane (166, 373). The immediate-early genes are transcribed immediately following entry and give rise to the regulatory proteins that impede host cell functions, such as gene expression, cell cycling, apoptosis, and the immune response, which allows the virus to take over the host cell’s DNA replication machinery (451). The immediate-early genes are under the control of a strong
promoter, major immediate-early promoter, which is the key regulator of CMV replication (284). Concurrently, the de-envelopment and disassembly of the virus particle allows the uncoated nucleocapsid and tegument proteins to be released into the cell and pass through the cytoplasm by latching onto microtubules (MT), allowing for rapid translocation to the nucleus (553). In cell types that are not permissive to CMV, the transcription of the major immediate-early promoter is repressed thus preventing viral replication (284).

**Figure 1-13: CMV Viral Replication Cycle** [Reprinted with permission from (553)]

CMV gene expression occurs in a temporal cascade with the immediate-early (α) proteins enabling the expression of early genes (β) that are required for DNA synthesis, and subsequently the late genes (γ) that encode virion structural and maturation proteins (451, 553). Upon entry into the nucleus, the linear viral DNA circularizes to act as a template for transcriptional activation of the origin of lytic replication (oriLyt), located in the UL region of the viral genome (25, 553). The oriLyt is a large structurally complex region (discussed in Section 3.5.2 Restriction Enzyme Digestion) that is transcriptionally transactivated by viral origin binding...
proteins, which initiates DNA replication (24, 88, 895). The CMV genome undergoes a rolling circle method of replication executed by various genes that code for polymerase, polymerase processivity factor, trimeric helicase-primase and single-stranded DNA binding protein, which is characteristic of herpesvirus DNA replication (24, 552) (reviewed in Chapter 3). DNA synthesis forms a long linear concatemer that lacks the genomic terminal fragments (530). During DNA replication, late genes are transcribed, many of which encode capsid proteins that are required for virus assembly and the formation of the nucleocapsid (24). The linear concatemer is inserted into the nucleocapsid and once the complete length of the genome is inserted, encapsidation proteins recognize the herpesvirus conserved pac1 and pac2 sites located in the repeat regions of the genome, which are subsequently cleaved allowing the viral DNA to be released and packaged into the capsid “DNA encapsidation” for nuclear egress (208, 295, 530, 553). During nuclear egress, the nucleocapsid undergoes primary envelopment and de-envelopment at the nuclear membrane before the nucleocapsid is released into the cytoplasm, where it is coated with tegument proteins (295). With the help of cytoplasmic microtubules, the tegumented nucleocapsid is shuttled through the cytoplasm for the final or “secondary” envelopment into cytoplasmic membranes, arising by vesicle transport from the endoplasmic reticulum (ER) through the Golgi body (GB), which forms an ER-Golgi intermediate compartment (ERGIC) (726). The mature progeny virions are transported via the ERGIC to the cell surface and released from the cell as intact infectious viral particles (VP) (553). Similarly, incomplete enveloped particles that bud from infected cells are termed dense bodies (DB) (553). Dense bodies are non-infectious virions that contain an envelope membrane densely filled with tegument proteins, however they lack a nucleocapsid and dsDNA viral genome (451).

HCMV establishes latency in a number of different tissues sites where viral gene expression is restricted by cellular mechanisms and controlled by cell-mediated immune response, both of which prevent reactivation and lytic replication (182, 652). HCMV reactivation sporadically occurs when the immune response is altered, such as by immunosuppression, stress, viral infections, or inflammation, leading to changes in the host cellular milieu (241, 382, 447, 664, 706). When cellular proteins or transcription regulators are altered, for example when bone marrow-derived myeloid cells are stimulated to differentiate into their mature phenotype, it enables them to activate the HCMV major immediate-early promoter, leading to immediate-early gene expression and subsequent viral reactivation and lytic replication (485, 584, 761, 819).
Upon reactivation of latent virus, infectious virions are released and shed at the mucosal surfaces (553), however the reactivated HCMV quickly triggers the surveilling immune system and the cells harboring lytic replication are efficiently cleared. Though under active immune surveillance, viral reactivation and shedding still occurs. In order to overcome the strong T cell response and maintain low level persistent infection in the host, CMV encodes a number of immune evasion genes that function to downregulate MHC class I, interfere with the antigen presentation pathway, act as immune homologues, or prevent target cell apoptosis (673). The immune evasion genes (described in more detail in Section 3.5.9 Immune Modulatory Genes) act to evade immune surveillance and prevent the host from clearing CMV infection, thus promoting lifelong infection.

1.3.3.3 Immune Response to Cytomegalovirus Infection

Murine CMV (MCMV) and HCMV studies have documented the role of the innate immune response in detecting HCMV infection through Toll-like receptors (TLRs), specifically TLR9, TLR3, and TLR2, which triggers the release of proinflammatory cytokines that serve to activate the adaptive immune response (76, 77, 165, 209, 812). Cytomegalovirus infection is highly immunogenic and elicits a cellular and humoral immune response from lymphocytes, predominantly T cells, though not exclusively, as NK cells and B cells are also required (674). The first line of cellular defense comes from NK cell-mediated cytotoxicity and lysis of CMV-infected cells, however this protection is only transient as NK cells decline in the later stages of infection (29, 71, 106, 195, 668, 844, 873). Interestingly, it has recently been shown that CMV-primed NK cells secreting IFN\(\gamma\) may control retroviral infection by increasing the number of activated anti-retroviral CD8\(^+\) T cells, which could also have implications for controlling HIV infection (269). The humoral immune response to immunogenic glycoproteins and phosphoproteins plays a role in preventing transmission and viral dissemination (397). Strain-specific neutralizing antibodies target the highly conserved envelope glycoproteins, particularly gB and gH, on the surface of the virion and have been shown to prevent CMV disease following vertical or organ transplant transmission of the virus (235, 435, 553, 815, 838). Non-neutralizing anti-CMV antibodies are also present in CMV seropositive individuals and are involved in reducing viral dissemination (397). Following primary infection, anti-HCMV IgM antibodies are transiently expressed followed by lifelong IgG antibodies, although IgG antibody titres decrease with age (283, 451). The benefit of the humoral immune response stems from evidence derived
from the Towne vaccine studies, as well as studies evaluating passive transfer of anti-CMV hyperimmunoglobulin or maternal transfer of anti-CMV IgG, which demonstrated that anti-CMV antibodies were involved in reducing CMV disease and may be associated with the prevention of congenital infection (588, 646, 648, 649, 772, 773).

CMV is predominantly controlled by a robust T cell-mediated immune response, with greater than 10% of the total T cell population being CMV-specific (232, 297, 810). The functional HCMV-specific T cell response greatly exceeds those specific for other viruses, including herpesviruses such as HSV and VZV (35, 119). CD8+ cytotoxic T lymphocytes (CTLs) are the dominant antiviral response that controls viral replication by lysing CMV-infected target cells, as shown both in vitro and through adoptive transfer studies in murine models and in bone marrow transplant (BMT) patients (89-91, 236, 689, 741, 860). Furthermore, CTLs have been shown to control CMV replication and limit CMV disease, as CMV-specific CTL responses were associated with survival following BMT receipt (668, 682, 683). Both CD4+ and CD8+ T cells play important roles in immune surveillance and mediating the anti-CMV immune response (660). With respect to CD4+ T cells, the aforementioned adoptive transfer studies also documented the importance of CMV-specific CD4+ helper T cells, which function to elicit and maintain the CTL and antibody responses over time (236, 281, 740, 860). Furthermore, CD8+ T cell depletion studies in mice have demonstrated that CD4+ T cells were able to maintain viral control and clear CMV from infected tissues (396).

During acute CMV infection, the memory T cell populations exhibit a classical central-memory phenotype (e.g. CD27+CD28+), however during latent CMV infection, CMV reactivation and antigen re-exposure stimulates the CMV-specific T cell memory population to differentiate into large pools of effector-memory T cells (e.g. CD27−CD28−), which are maintained with high frequency at mucosal sites (Figure 1-14) (28, 47, 148, 282, 367, 449, 553, 566, 601). Furthermore, an age-related increase in the magnitude of CMV-specific functional (secreting IFNγ and/or TNFα) effector-memory T cells, known as “memory inflation” was observed when PBMCs from donors of various age groups were stimulated with CMV peptides (426, 427, 845). Memory inflation occurs with certain persistent viral infections, such as CMV, and refers to the age-related expansion of functional virus-specific memory T cells (412, 601), which is in sharp contrast to other HIV viral vectors being evaluated, namely adenovirus or VZV, in which virus-specific T cell responses decrease with age (35, 739). Unlike during chronic
viral infections, such as HIV, in which effector-memory cells display an exhausted phenotype (49, 199), chronic CMV infection is associated with low level antigen exposure, which induces effector-memory CD8$^+$ cells that display an “inflation” phenotype (CD27$^-$/CD28$^-$/PD-1$^-$) and retain their functional antiviral capacity supported by CD4$^+$ T cell help (352, 770, 771, 861).

Figure 1-14 CMV-Specific Memory T cell Phenotypes [Reprinted from Trends in Immunology, (601), Copyright 2012, with permission from Elsevier]

1.3.4 Overarching Hypothesis

In view of the described features of CMV, we believe that CMV as a replicating vaccine vector has the potential to elicit substantial mucosal-oriented, persistent, and functional immune responses. Thus, we hypothesize that using a chronic reactivating virus as an HIV vaccine vector will provide durable and persistent in situ mucosal T cell effector memory responses by eliciting protective effectors at the portal of HIV entry and in a timeframe most appropriate to inducing clinically relevant protection. In addition, we predict that we will be able to evaluate the utility and effectiveness of CMV as a viral vector in a cynomolgus macaque non-human primate/SIV challenge model.
1.4 Non-Human Primate Models

To evaluate CMV as an HIV/SIV viral vector, a non-human primate model will be utilized. Animal models, particularly non-human primates (NHPs), have been fundamental for the study of HIV with respect to prevention, pathogenesis, treatment, and vaccination (245, 342). The primary model systems used in HIV research are certain NHPs (293, 579, 596) however, murine models such as bone marrow-liver-thymus (BLT) humanized mice are showing increasing utility (94, 215, 230, 434, 568, 804). With the close physiological and immunological similarities to humans, NHP models represent a practical model to assess the efficacy of an HIV/SIV vaccine as SIV infection in Old World macaques closely recapitulates HIV disease progression in humans, albeit at a faster rate (192). NHP models are important for the optimization of the vaccination regimen and challenge phase, with regards to the number and route of immunizations and challenges, dosage, and allows for the control of the strain of SIV used for challenge, thus enabling heterologous challenge. Furthermore, these models allow for synchronous infections, and the ability to perform repeated sampling in order to closely examine pathogenesis and the immune response during the acute stages of infection. Following necropsy, an abundance of information can be gathered through access to various tissue samples, which can better inform research outcomes and direct future vaccine development. This information includes, but is not limited to, an invaluable understanding of possible immune correlates of protection, tissue-specific viral dissemination, and the evaluation of latent viral reservoirs.

With respect to NHPs, chimpanzees (*Pan troglodytes*) are the closest models to humans with greater than 96% genome homology (172). Chimpanzees have been used for HIV vaccine research as they are one of the only NHPs that can be infected with HIV (19, 63, 65, 93, 238, 277, 299, 662), however the pathogenic response was limited and the animals rarely progress to AIDS (370, 579). Furthermore, chimpanzees are extremely costly and more importantly have become a protected species due their endangered species status, thus limiting their use in biomedical research (370, 579). Initial attempts were made to infect pig-tailed macaques (*Macaca nemestrina*) with HIV, however only transient infection was observed (9, 275). Alternatively, Old World monkeys, namely cynomolgus macaques (*Macaca fascicularis*), rhesus macaques (*Macaca mulatta*), African green monkeys (*Chlorocebus aethiops*), and sooty mangabeys (*Cercocebus atys*) are widely utilized SIV animal models, as SIV infection in these macaques more closely mimics HIV infection in humans. These four NHP species can be
differentiated by whether they are natural or non-natural SIV hosts and their associated pathogenic features (Table 1-2). Like humans, cynomolgous macaques (CMs) and rhesus macaques (RMs) are non-natural hosts, thus are typically used for examining treatment approaches and vaccine development. The clinical outcome in SIV-infected CMs and RMs is similar to HIV-infected humans, as marked by a high SIV viral load, a decline in CD4+ T cells, microbial translocation and chronic immune activation leading to the progression to simian AIDS (SAIDS) (125). Conversely, African green monkeys (AGMs) and sooty mangabeys (SMs) are SIV natural hosts and provide valuable information for comparative biology and pathogenesis, however these macaques models are not useful in a vaccine setting, as they do not exhibit typical SIV pathogenesis or progress to SAIDS, despite having high levels of replicating virus (125). As such, the focus from this stage forward will be placed on CMs and RMs.

Table 1-2 Natural versus Non-Natural SIV Hosts [Adapted from (125). Reprinted with permission from AAAS]

<table>
<thead>
<tr>
<th>Natural Hosts (AGM, SM)</th>
<th>Phenotype</th>
<th>Non-Natural Hosts (CM, RM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>SAIDS</td>
<td>Yes</td>
</tr>
<tr>
<td>Healthy</td>
<td>Level of peripheral CD4+ T cells</td>
<td>Low</td>
</tr>
<tr>
<td>High</td>
<td>Viral load</td>
<td>High</td>
</tr>
<tr>
<td>Yes</td>
<td>Virus cytopathicity</td>
<td>Yes</td>
</tr>
<tr>
<td>Ineffective</td>
<td>Host immune control</td>
<td>Ineffective</td>
</tr>
<tr>
<td>Yes, stable</td>
<td>Depletion of mucosal CD4+ T cells</td>
<td>Yes, progressive</td>
</tr>
<tr>
<td>No</td>
<td>Mucosal immune dysfunction/microbial translocation</td>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
<td>Chronic immune activation</td>
<td>Yes</td>
</tr>
<tr>
<td>Tem &gt; Tcm</td>
<td>Pattern of infected cells</td>
<td>Tcm &gt; Tem</td>
</tr>
<tr>
<td>Rare</td>
<td>Vertical transmission</td>
<td>Frequent</td>
</tr>
</tbody>
</table>

A wide variety of reagents are available for Indian rhesus macaques, however they are becoming increasingly difficult to use for HIV research as they are in high demand. Furthermore,
in 1978 the United States placed a ban on the import of Indian rhesus macaques, thus forcing researchers to rely solely on established breeding colonies, which have since been struggling to keep up with the demand (158). Efforts are underway to develop and/or improve large breeding programs for Indian rhesus macaques, however many researchers have begun using Chinese rhesus macaques and cynomolgus macaques, which are more readily available (676). It has been well established that the duration of SIV pathogenesis and progression to SAIDS is shorter in Old World macaques (~0.5-3 years) than in treatment-naïve humans (~8-10 years) (342, 370). However, in comparison to rhesus macaques, cynomolgus macaques survive longer following SIV infection and thus more closely recapitulate HIV disease progression in humans (596, 676). It has been proposed that since SIVmac was passaged through Indian rhesus macaques, the virus may be more adapted to rhesus macaques and thus exhibits an attenuated pathogenicity in cynomolgus macaques (596, 676). Compared to Indian rhesus macaques, SIVmac251-infected cynomolgus macaques of Mauritius origin demonstrate a lower steady state SIV plasma viral load with a less drastic CD4+ T cell loss and a higher frequency of virus-specific immune responses, resulting in a slower progression to SAIDS (676, 822). These clinical outcomes can be both advantageous and disadvantageous for evaluating SIV vaccine efficacy in a cynomolgus macaque model. A less pathogenic infection may provide a more accurate estimate of vaccine efficacy in humans; however it may also make it difficult to observe subtle vaccine-induced reductions in set-point viral load, though this limitation may be overcome by ensuring a high enough statistical power is achieved in order to observe vaccine-related differences between groups (596, 676). Taken together, the reduced pathogenicity and delayed progression to SAIDS observed in SIV-infected cynomolgus macaques allows for a longer period to monitor vaccine efficacy, during which time extensive tissue sampling can be performed in an effort to elucidate potential immune correlates of protection. For this reason and others highlighted below, our laboratory has chosen to evaluate and develop HIV/SIV vaccine candidates in a cynomolgus macaque model.

Cynomolgus macaques (known as long-tailed or crab-eating macaques) are a widely utilized non-human primate model for preclinical biomedical research (596), including infectious disease research (118, 303, 455, 839), transplant research (259, 419, 734), and are becoming an increasingly popular NHP model for HIV research (50, 108, 261, 596, 833, 878, 884). The complete cynomolgus macaque genome has recently been sequenced and characterized for
cynomolgus macaques derived from Mauritian (234), Vietnamese (897), and Malaysian (355) geographic origin. Approximately 500 years ago, a group of cynomolgus macaques diverged from Indonesia aboard a ship that made a stopover on the island of Mauritius, resulting in the formation of a small founder population on the island (456, 806, 807, 830). A population expansion of these founder animals has lead to a relatively genetically homogenous population of Mauritian cynomolgus macaques, with high genetic variation from cynomolgus macaques of other geographic origins (408). Rhesus macaque genetics have been more thoroughly characterized, though much effort is now being spent on delineating the genetic composition of cynomolgus macaques (408, 728, 884), which is especially true for cynomolgus macaques of Mauritius origin (MCMs). MHC class I (110) and class II alleles (181, 223, 460, 464, 473, 474, 599) have been described for cynomolgus macaques of various geographic origins, though Mauritian cynomolgus macaques have more homogenous MHC alleles with only seven defined haplotypes (105, 600, 884). Certain MHC class I haplotypes have been shown to confer protection from HIV infection in humans (HLA-B27, HLA-B57) (547, 713), or from SIVmac infection in rhesus macaques (Mamu-A*01, Mamu-A*1303, Mamu-B*08, Mamu-B*17) (478, 479, 563, 564, 598, 901). Similarly, both protective from infection (H3, H6) and susceptibility to infection (H2, H5) haplotypes have been identified in Mauritian cynomolgus macaques (261). In vaccination studies, the low MHC diversity and genetic homogeneity observed in cynomolgus macaques is advantageous as it allows for a more controlled study population, MHC-matched animals, less variables and genetic contributions that may hinder the outcome of the vaccine. Furthermore, this limited genetic diversity may assist in elucidating the correlates of protection as it can simplify the identification of cell-mediated immune responses.

In addition, historically cynomologus macaques have been more readily accessible in Canada. At the time of the studies outlined in this thesis, a well-established cynomolgus macaques breeding colony was in operation at Health Canada in Ottawa, Canada; thus allowing for ease of access, as well as the opportunity to collaborate with the National HIV and Retrovirology Laboratory at the Public Health Agency of Canada. Researchers, veterinary and animal care staff from both agencies had considerable experience working with cynomolgus macaques and an expertise in the clinical, psychosocial, and behavioural aspects of these animals. The colony was formed in 1983 and was in operation for 28 years before closing in 2011. At the time of closure, the colony consisted of 189 cynomolgus macaques, which
comprised 85 males and 104 females, 129 of which were colony-bred and 60 animals were imported. The origin of the animals varied between Mauritius background (N=60), Indonesian/Filipino (N=73), or Mauritian/Indonesian mixed origin (N=56). The animals ranged from 1 to 25 years of age and the majority of animals were adults (12-19 years old). Another advantage for using cynomolgus macaques as a model for HIV/SIV vaccine development arises from a previous study conducted by our group, which demonstrates that cynomolgus macaques are permissive to productive infection with human VZV (877), the other herpesvirus being examined by the MacDonald laboratory. Unlike in rhesus macaques, in which an attenuated strain of human VZV was shown to have limited infectivity, replicative capacity, and immunogenicity (788), cynomolgus macaques inoculated with a non-attenuated human VZV displayed VZV-specific cellular and humoral immune responses comparable to what has been observed in humans (877). Thus, with the long-term goal of comparing and contrasting CMV and VZV as viral vectors with the potential for dual vaccination, we concurrently began developing CMV in a cynomolgus macaque model.

NHP models provide invaluable information regarding the optimization, efficacy, and safety of pre-clinical HIV vaccine candidates. However, it is also imperative to recognize that the information gained in NHP studies does not always translate into a human setting. Although comprised of largely the same genes, the complete SIVmac genome is only 40% homologous with HIV-1 (126). Furthermore, NHP models are in fact only models for humans and it is clear that macaques have a different genetic composition, are exposed to different pathogens, and have different life events, all of which may impact the immune response. Unfortunately, this was well demonstrated in the Step trial in which the safety and efficacy trials performed in rhesus macaques were not predictive of the deleterious outcomes observed in human clinical trials (as described in Sections 1.2.1 Obstacles in the Development of an HIV Vaccine and 1.2.3.2 Step Trial). It was clear from the Step trial that pre-existing Ad5 immunity impacted HIV acquisition in the vaccinated group, thus it will be important moving forward to examine anti-vector immune responses in individuals prior to vaccination with a viral vector. Our study has the advantage of evaluating pre-existing immunity and CMV vector-mediated responses in cynomolgus macaques since our previous assessment of CMV seroprevalence in the cynomolgus macaque colony at the Public Health Agency in Canada was close to 100% (22). This highlights the need for evaluating HIV/SIV vaccine candidates in multiple NHP models in an effort to detect any potential adverse
effects prior to moving forward into human clinical trial. All things considered, it is the responsibility of the researchers to understand the confines of NHP models when interpreting the outcomes from SIV vaccine trials, and it is essential to take into account such limitations when determining whether the NHP trials warrant evaluation in human studies. In order to ensure safety and avoid repeating previous failures, a strong focus must be placed on how restrictions conferred by NHP models will impact translation into human HIV vaccine clinical trial.

### 1.5 Thesis Background and Objectives

Upon commencing my PhD in 2009, a cynomolgus macaque CMV infectivity trial had just been initiated, and was primarily designed and led by Dr. Aruna Ambagala, a post-doctoral fellow in the MacDonald laboratory. Throughout the trial, I was involved in sample processing and screening, in addition to designing and performing a number of outstanding experiments once the trial was complete. After Dr. Ambagala’s tenure in the MacDonald laboratory, I took on the responsibility of compiling the data, designing and performing outstanding experiments, as well as writing the manuscript in a continued effort to advance our focus of evaluating CMV-based HIV/SIV vaccines. The results of this infectivity trial provided the basis for my thesis project and thus will be briefly described herein.

#### 1.5.1 Species-Specificity of Non-Human Primate Cytomegaloviruses


It has been previously documented that cynomolgus macaques display evidence of CMV infection with the majority of captive cynomolgus macaques being CMV-seropositive, however cynomolgus macaque host-specific virus has not been isolated or genetically characterized (603, 604). Thus, to initiate the evaluation of a CMV-based HIV viral vector in the cynomolgus macaque model, we began by testing a more readily available virus, rhesus macaque cytomegalovirus (RhCMV), due to the close phylogenetic relationship between rhesus and cynomolgus macaques. In addition to accessible reagents, RhCMV had been genetically
characterized and cloned as a bacterial artificial chromosome (BAC), making it amenable to manipulations required for our evaluation. We constructed a recombinant replication-competent RhCMV constitutively expressing enhanced green fluorescence protein (eGFP) to enable differentiation between productive RhCMV-eGFP infection and endogenous cynomolgus macaque CMV. It has been well established that cytomegaloviruses have co-evolved with their host species following co-speciation over 80 million years ago, resulting in highly species-specific viruses (524). However, it has been shown that certain non-human primate CMV strains can overcome the species-specific barrier in vitro (75, 470), and the extent of in vivo host-restriction has not been well documented (36, 553). Thus, whether cross-species infection can occur between closely related species, such as Old World monkeys (OWM), remains to be elucidated. However, given the close species homology between rhesus macaques and cynomolgus macaques, we hypothesized that rhesus macaque CMV (RhCMV) would be able to infect cynomolgus macaques. In this study, we inoculated cynomolgus macaques with RhCMV-eGFP in an effort to establish productive infection, marked by evidence of RhCMV-eGFP-specific immune responses and detectable RhCMV-eGFP in the tissues and/or being shed in the urine (Figure 1-15).
Figure 1-15 Inoculation and Sampling Schedule. Twelve cynomolgus macaques were randomly assigned into three groups, RhCMV-eGFP (N=6), UV-inactivated RhCMV-eGFP control (N=2), and media control (N=4). The animals received one subcutaneous inoculation at week 0 with $7 \times 10^7$ PFU of RhCMV-eGFP or UV-inactivated RhCMV-eGFP, or media alone. A subset of animals was subcutaneously boosted at week 8 with $2 \times 10^7$ PFU of RhCMV-eGFP or UV-inactivated RhCMV-eGFP, while the remaining animals received media alone. At week 23, one UV-inactivated RhCMV-eGFP control animal (C09-001M) and three media control animals (C09-002M, C09-006M) received an intravenous inoculation with RhCMV-eGFP ($7 \times 10^7$ PFU). The sample collection schedule is described.

A variety of viral detection and immunology assays were performed to extensively examine numerous tissues (blood, urine, BAL, saliva, lymph nodes, oral and genital swabs) over the course of the trial (Figure 1-15), and for a subset of animals the follow-up was extended to 4 years. At the commencement of the study, all animals were clinically healthy with comparable body weights and immunophenotyping profiles. Following the inoculations, none of the animals displayed any clinical symptoms of CMV infection or body weight changes throughout the course of the study. To monitor the general health status of the animals and enumerate lymphocyte populations, complete blood counts comprising 20 haematology markers, and immunophenotyping were performed. Cellular immune responses to eGFP and a CMV-specific gene (IE-1) were examined by intracellular cytokine staining on PBMCs isolated at baseline (week -1) and following the three inoculations (weeks 6, 13, and 29). To determine if eGFP-specific antibodies were present in the serum of these animals, an eGFP-specific ELISA was performed. With respect to virologic assays, samples from easily accessible sites of CMV infection and routine shedding (including urine, BAL, saliva, oral and genital swabs, lymphocytes) were cultured on fibroblast cells to screen for concurrent eGFP expression and CMV cytopathic effect (CPE). Furthermore, total protein isolated from urine co-cultures displaying CMV CPE, was probed with anti-eGFP antibody to determine if eGFP could be detected by Western blot. An exhaustive number (3,233 total) of real-time quantitative polymerase chain reaction (qPCR) assays were performed in whole blood, plasma, urine (whole urine, urine pellet, and DNA-derived from urine co-cultures), BAL, saliva, and lymph node biopsies in an attempt to amplify RhCMV-eGFP-specific products.

None of the described experiments showed any evidence for the presence of eGFP or RhCMV-specific genes, and/or immune responses to these genes. Thus, the result of this extensive analysis strongly suggests that RhCMV-eGFP was not able to infect cynomolgus macaques. In this setting, the species-specific barrier could not be broken between these Old
World macaques and it is clear that the study of CMV is restricted to its host species. Thus, in an effort to develop and evaluate CMV as a viral vector for an HIV/SIV vaccine in a cynomolgus macaque model, host-specific CMV will be required. To facilitate this, we are the first group to isolate and characterize a novel cynomologus macaque CMV (CyCMV).

1.5.2 Thesis Objectives

Overview: This thesis will describe my work focused on isolating a novel cynomolgus macaque CMV (CyCMV), as well as the confirmation and characterization of this virus as a cytomegalovirus. Furthermore, I performed next-generation sequencing, open reading frame assignment, and genomic analysis of the viral genome, which will provide invaluable information for targeted insertions of vaccine antigens, attenuation mutations, and the like, in an effort to develop CyCMV as an HIV/SIV vaccine vector. The sequenced CyCMV genome enabled targeted cloning of the large linear CyCMV genome as a Bacterial Artificial Chromosome, in preparation for the insertion of the vaccine antigens, as well as to mediate viral attenuation and safety manipulations. Ultimately, the goal of this thesis project is to develop CyCMV as a viral vector to be utilized in the generation of CyCMV-SIV vaccine constructs for evaluation in a cynomolgus macaque SIV challenge model.

Main Objective: To isolate, genetically characterize, and develop cynomolgus macaque cytomegalovirus (CyCMV) as an HIV/SIV viral vector for evaluation in a non-human primate model.

Specific Objectives

1. Isolation and characterization of cytomegalovirus from cynomolgus macaques (Chapter 2)

2. Identification and preliminary classification of the isolated virus as a novel CMV (Chapter 2)

3. Next-generation DNA sequencing of CyCMV (Chapter 3)

4. Bioinformatic assembly of the CyCMV DNA sequence reads (Chapter 3)

5. Genomic Analysis of CyCMV (Chapter 3)

6. Cloning of CyCMV as a Bacterial Artificial Chromosome (Chapter 4)
Chapter 2

2 Isolation and Characterization of Cynomolgus Macaque Cytomegalovirus (CyCMV)

Angie K. Marsh¹, Aruna P. Ambagala², Jacqueline K. Chan², Richard Pilon³, Jocelyn Fournier⁴, Paul Sandstrom³, David O. Willer¹, Kelly S. MacDonald¹,²

¹Department of Immunology, University of Toronto, ON, Canada, M5S 1A8, ²Department of Microbiology, Mount Sinai Hospital, ON, Canada, M5G 1X5, ³National HIV & Retrovirology Laboratories, National Microbiology Laboratory, Public Health Agency of Canada, Ottawa, ON, Canada, K1A 0L2 ⁴Scientific Services Division, Health Products & Food Branch, Health Canada, Ottawa, ON, Canada K1A 0L2

Author Contributions: All experiments and figures described in this chapter were performed and generated by myself (A.K.M.), with exception of a portion of the Western blot experiment, which was performed by J.K.C. For the Western blot experiment (Section 2.5.6; Figure 2-6), I performed the viral infections, acquired the antibodies, and prepared the lysates, while J.K.C. conducted the Western blot assay. A.K.M., A.P.A., D.O.W. and K.S.M. were involved with conceiving, designing, and analyzing the experiments. R.P., J.F., P.S. provided the cynomolgus macaque urine samples from our collaborative studies being conducted at the non-human primate colony in Ottawa. Furthermore, the transmission electron microscopy was performed by Battista Calvieri and Steven Doyle at the Microscopy Imaging Facility at the University of Toronto (whom are acknowledged in the Acknowledgements section at the end of this chapter).

Part of the work contained in this Chapter is from the following publication, and only the work primarily conducted by Angie Marsh was included: © Springer and Archives of Virology, 158 (5), 2013, 955-965, Establishment of an immortal cynomolgus macaque fibroblast cell line for propagation of cynomolgus macaque cytomegalovirus (CyCMV). Ambagala, A. P. N., Marsh, A. K., Chan, J. K., Mason, R., Pilon, R., Fournier, J., Sandstrom, P., Willer, D. O., MacDonald, K.S., original copyright notice is given to the publication in which the material was originally published, by adding; with kind permission from Springer Science and Business Media.
2.1 Chapter Overview

Our recent efforts evaluating cross-species infection of RhCMV in cynomolgus macaques have outlined the importance of using host-specific CMV for *in vivo* studies. In this regard, I isolated cynomolgus macaque-specific CMV (CyCMV; Mauritius strain) and confirmed the virus exhibited the characteristic phenotypic and genomic features for classification as a cytomegalovirus. The isolation and characterization of CyCMV is described in detail in this chapter.

**Main Objective:** To isolate, genetically characterize, and develop cynomolgus macaque cytomegalovirus (CyCMV) as an HIV/SIV viral vector for evaluation in a non-human primate model.

**Specific Objectives**

1. Isolation and characterization of cytomegalovirus from cynomolgus macaques

2. Identification and preliminary classification of the isolated virus as a novel CMV

3. Next-generation DNA sequencing of CyCMV (Chapter 3)

4. Bioinformatic assembly of the CyCMV DNA sequence reads (Chapter 3)

5. Genomic Analysis of CyCMV (Chapter 3)

6. Cloning of CyCMV as a Bacterial Artificial Chromosome (Chapter 4)
2.2 Abstract

Cytomegalovirus (CMV) is a highly species-specific virus that has co-evolved with its host over millions of years and thus restricting cross-species infection. Furthermore, an evaluation of species-specificity between closely related macaques revealed that rhesus macaque CMV (RhCMV) cannot productively infect cynomolgus macaques. Thus, in an effort to develop and evaluate CMV as a viral vector for an HIV/SIV vaccine in the cynomolgus macaque model, host-specific CMV will be required. To facilitate this, cynomolgus macaque cytomegalovirus (CyCMV) was isolated from the urine of cynomolgus macaques of Mauritius origin. CyCMV Mauritius strain was similar in size and ultrastructural morphology to previously characterized CMVs, and exhibited characteristic CMV-specific cytopathic effect and viral growth properties in tissue culture. Phylogenetic examination of two highly conserved CMV genes, DNA polymerase and glycoprotein B, demonstrated that CyCMV clusters with other Old World macaques, however sequence analysis revealed a unique difference between CyCMV strains with respect to the glycoprotein B coding region, with CyCMV Mauritius strain being more closely related to RhCMV than to the other characterized CyCMV Ottawa strain, which was isolated from cynomolgus macaques of a different geographic origin. CyCMV productively replicates in human, rhesus, and cynomolgus macaque fibroblast cell lines and classically downregulates MHC class I protein expression in infected cells. In addition to establishing a novel model to evaluate CMV-based HIV/SIV vaccines, CyCMV may serve as an additional non-human primate model to study CMV biology, pathogenesis, and HCMV vaccine development.
2.3 Introduction

Cytomegalovirus (CMV) is a large double-stranded DNA herpesvirus categorized in the beta herpesvirus family. CMV has a broad tissue tropism, primarily replicating within mucosal sites, and infecting a number of different cell types, including cells of myeloid origin, fibroblast cells, epithelial cells, endothelial cells, neuronal cells, smooth muscle cells, and hepatocytes (553, 764, 765). Human CMV (HCMV) is a ubiquitous pathogen with a global seroprevalence of 45-100% depending on the geographic region, and is transmitted via genital secretions, breast milk, saliva, urine, placental transfer, or through transplantation (114, 553, 630). HCMV establishes lifelong infection and elicits a robust CMV-specific T cell response with greater than 10% of the total T cell population designated to controlling CMV infection, which is typically asymptomatic in immunocompetent hosts (232, 297, 810). However, in immunocompromised individuals, such as neonates, transplant patients, individuals receiving chemotherapy, or AIDS patients, the immune system is unable to control HCMV infection, which can cause severe disease including congenital CMV disease, retinitis, pneumonitis, hepatitis, esophagitis, gastritis, and enterocolitis (82, 283, 766).

In addition to HCMV, host-specific CMVs have been isolated from a variety of different species including mice (671), rats (846), guinea pigs (405, 731), and tree shrews (42), as well as non-human primates, such as chimpanzees (198, 405), baboons (75), colobus guereza (658), and rhesus macaques (337, 691). Another widely utilized non-human primate model for infectious disease and transplant research is cynomolgus macaques, however a cynomolgus macaque CMV has not been characterized. With our primary focus of evaluating CMV as a viral vector for an HIV/SIV vaccine in a cynomolgus macaque model, we recently examined whether the previously characterized and most widely used non-human primate CMV, rhesus macaque CMV (RhCMV) (337) could establish cross-species infection in closely related cynomolgus macaques. Our results strongly suggested that RhCMV was not able to infect cynomolgus macaques and the species-specific barrier could not be broken between these closely related non-human primates (Marsh, A. K., Ambagala, A. P., et al. submitted) (previously reviewed in Section 1.5.1 Species-Specificity of Non-Human Primate Cytomegaloviruses). Thus, in a continued effort to develop and evaluate CMV-based HIV/SIV vaccines in cynomolgus macaques, I have isolated and characterized host-specific cynomolgus macaque CMV (CyCMV). We concurrently isolated
CyCMV from cynomolgus macaques of two distinct geographic origins, including the Ottawa
strain from macaques of Indonesian/Filipino origin (22), and the Mauritius strain from macaques
of Mauritius origin (described herein) to evaluate potential viral evolutionary differences.

2.4 Material and Methods

2.4.1 CyCMV Isolation

Urine samples were collected from cynomolgus macaques (*Macaca fascicularis*) of
Mauritius origin using negative pressure bladder catheterization or from the cage pan in which
the urine was filtered through a 0.45 µm filter. The samples were centrifuged at 900 x g for 30
mins at 4°C after which the supernatants were collected and mixed 1:1 with 2X MEM
(supplemented with 2X antibiotic-antimycotic and 20 mg/ml gentamycin; Invitrogen). The cell
pellets were resuspended with 1 ml of Dulbecco’s Modified Eagle’s Medium (DMEM; Sigma)
supplemented with antibiotic-antimycotic, and 10 mg/ml gentamycin (Invitrogen). Centrifuged
urine samples were further clarified by ultracentrifugation at 20,000 x g for 30 mins at 4°C and
the pellets were resuspended in 500 ul of DMEM (supplemented with antibiotic-antimycotic and
10 mg/ml gentamycin). Human fetal lung fibroblast (MRC-5) cells (381) were obtained from the
American Type Culture Collection (ATCC) and seeded 1:2 in a 12-well tissue culture plate 2
days prior to inoculation. The samples were each plated in triplicate wells of MRC-5 cells,
spinoculated at 2,000 x g for 30 mins at 4°C, and incubated for 2-3hrs at 37°C in a 5% CO2
incubator. Following the incubation, the inoculum was aspirated and 2 mls of DMEM +
supplemented with 10% FBS (Wisent Bioproducts), 100 U/ml penicillin, and 100 µg/ml
streptomycin (Sigma) was added. The medium was changed the following day and monitored
daily for CMV cytopathic effect (CPE). After approximately one month in culture, the
monolayers reached 100% CPE and CyCMV (Mauritius strain) was propagated at a 1:2 split
ratio on MRC-5 cells.

2.4.2 Viral DNA Isolation

Fifteen 75cm² tissue culture flasks containing CyCMV-infected MRC-5 cells were used
for viral DNA isolation. The infected monolayer was used for intracellular isolation and the
supernatant was harvested for extracellular isolation. The monolayer was washed once with PBS,
and viral DNA was extracted with 1.5 ml of HIRT extraction buffer (2X TE and 1.2% SDS)
for 20 mins at room temperature. The lysate was collected, treated with 1M sodium chloride, and placed at 4°C overnight. The lysate was centrifuged at 27,000 x g for 35 mins at 4°C and the supernatant was collected. Concurrently, the harvested supernatant containing extracellular virus was centrifuged at 2,000 x g for 10 mins at room temperature to pellet any remaining cells and filtered with a 0.45 µm filter. The supernatant was then centrifuged at 71,000 x g for 1 hr at 4°C and the pellet was dissolved in TE. The intracellular and extracellular samples were separately treated with RNAse (60 µg/ml RNAse A and 160 U/ml RNAse T1; Fermentas) for 2 hrs at 37°C followed by a Pronase treatment (1mg/ml; Roche) for 2 hrs at 37°C. Residual protein contamination was removed with one phenol and two phenol chloroform extractions, and the viral DNA was precipitated with 0.3M sodium acetate and two volumes of absolute ethanol at -20°C overnight. Viral DNA was clarified by centrifugation at 17,000 x g for 30 mins at 4°C and dissolved in TE. To further purify the isolated viral DNA, the samples were overlaid on a discontinuous 5-20% sucrose gradient containing ethidium bromide (2ug/ml) and centrifuged at 200,000 x g for 2.5 hrs at 4°C. Separated viral DNA was visualized by UV-illumination, harvested, and diluted in 1.5 volumes of water, 0.3 M sodium acetate, and 2.5 volumes of absolute ethanol overnight at -20°C. The samples were centrifuged at 15,000 x g for 30 mins at 4°C, washed once with 70% ethanol, and the pellet was dissolved in water. To confirm the purity of the viral DNA isolation, 1 µg of isolated extracellular viral DNA was digested with 20U of BamHI or HindIII restriction enzymes (New England Biolab) and fractionated by gel electrophoresis on a 0.8% agarose gel.

2.4.3 Transmission Electron Microscopy

CyCMV-infected cells were prepared for electron microscopy as previously described (22). Briefly, CyCMV (Mauritius strain)-infected MRC-5 cells exhibiting 80% CPE were harvested by scraping the monolayer with a rubber cell scraper. Cells and supernatant were centrifuged at 3,000 x g for 10 mins at room temperature, the pellet was washed once with PBS, and pelleted again. Primary fixing solution (4% paraformaldehyde and 1% glutaraldehyde in 0.1 M phosphate buffer pH 7.2) was added to the pellet for 1 hr at room temperature and the pellet was then washed three times with 0.1 M phosphate buffer. The pellet was fixed in secondary fixing solution (1% osmium tetroxide in 0.1 M phosphate buffer pH7.2) for 1 hr at room temperature, and washed three times in 0.1 M phosphate buffer. The pellet was dehydrated in increasing ethanol concentrations, embedded with Epon resin, and cut into ultra thin sections for
staining. The sample was sequentially stained with uranyl acetate, followed by lead citrate, and examined using a HITACHI H-7000 transmission electron microscope by the Microscopy Imaging Facility at the University of Toronto.

2.4.4 Polymerase Chain Reaction Amplification

Total DNA was isolated from CyCMV-infected MRC-5 cells, derived from urine co-cultures, using a DNeasy Blood & Tissue Kit (Qiagen). CyCMV glycoprotein B (gB) was amplified using degenerate primer sequences (22), forward primer 5’-CTGATCGTA(A/G)GTGTGAAA-3’ and reverse primer 5’-TATCTCCTCCTCTA(T/C)AGGCATCTG-3’. Primers were designed to amplify full-length DNA polymerase (Pol) gene by designing primers upstream and downstream of the Pol ORF (504), forward primer 5’-CCCTGACCCTCAGCAGCGTG-3’ and reverse primer 5’-AGCAGCTCCTCGCAACCAAGG-3’. The PCR conditions were as follows: 95°C for 7 mins, 35 cycles of 95°C for 30 secs, 65°C (gB) or 55°C (Pol) for 30 secs, 72°C for 55 secs and one cycle of 72°C for 7 mins. PCR amplicons were cloned into the pCR-Blunt II-TOPO vector (Invitrogen) and transformed into chemically competent cells (Invitrogen). The full-length gene inserts were confirmed by Sanger sequencing, sequences were aligned by ClustalW alignment, and shaded using Geneious Pro 5.4.6 (229).

2.4.5 CyCMV Growth Kinetics

To compare the replication kinetics of CyCMV (Mauritius strain) in fibroblast cells from different species, cynomolgus macaque (MSF-T) (23), rhesus macaque (Telo-RF) (432), and human (MRC-5) (381) fibroblast cells were grown in 6-well tissue culture plates. The confluent monolayers were inoculated (in duplicates) with 500 ul of CyCMV at two different multiplicities of infections (MOI) of 0.3 and 0.03. After 2 hrs of adsorption at 37°C, 1 ml of fresh cell culture medium (DMEM + 10% FBS) was added to each well. The cells and the culture supernatant were harvested on days 2, 4, 6, 8 and 10 post-infection and frozen at -80°C with 0.2 M sucrose added. The viral titers of each sample were determined by standard plaque assay on MRC-5 cells. Briefly, tenfold serial dilutions of each sample were prepared and used to inoculate confluent monolayers of MRC-5 cells in 6-well plates. After 2 hrs of adsorption at 37°C, the inoculum was removed and the cell monolayer was overlaid with a 1:1 mixture of 2X MEM (containing 10 % FBS and 4 mM L-glutamine; Quality Biologicals Inc.) and 2% low-melting
agarose (type VII; Sigma). The plates were incubated at 37°C for 12-14 days, and the plaques at the end-point dilutions were counted.

2.4.6 Western Blot

Mock- or CyCMV-infected MSF-T, Telo-RF and MRC-5 cells were lysed in RIPA buffer (50 mM Tris HCl, pH 8, 150 mM NaCl, 1 % NP-40, 0.5 % sodium deoxycholate, 0.1 % SDS, and 1 mM EDTA, adjusted to pH 7.4). Total clarified protein (10 µg) from each sample was resolved on a 4-20% TGX gel (Bio-Rad) and transferred to a nitrocellulose membrane (GE). The membrane was probed with HC-10 (HLA heavy chain) kindly provided by Dr. David Williams at the University of Toronto (786), rabbit anti-b2 microglobulin (M8523, Sigma), or mouse anti-actin (Sigma) antibodies, followed by the relevant secondary antibodies (anti-mouse-HRP or anti-rabbit-HRP, Cedarlane). The proteins were detected using a Luminata Western HRP Kit (Millipore) according to the manufacturer’s instructions.

2.5 Results

2.5.1 Isolation of CMV from Mauritius Cynomolgus Macaques

Cynomolgus macaque CMV (CyCMV) Mauritius strain was isolated from urine cocultures of healthy cynomolgus macaques of Mauritius origin. The primary isolates were propagated on human fetal lung fibroblast (MRC-5) cells, as previous work from our group demonstrated that the slow replication kinetics of MRC-5 cells best supports the slow viral replication of primary isolates (22). CyCMV displayed characteristic CMV growth properties and the primary isolates were cultured for 4 weeks before 100% CMV-specific cytopathic effect (CPE) was observed, characterized by the presence of enlarged rounded cells in the form of clusters. The cultures were passaged with a 1:2 split ratio in an effort to propagate the virus for downstream viral DNA isolation and further characterization. The two most rapidly growing isolates (C09-007M and C09-012M) were selected for large-scale in vitro propagation, which achieved approximately 10 passages before the viral DNA was isolated. Thus, CyCMV Mauritius strain would be classified as a low-passaged virus compared to the previously isolated CyCMV Ottawa strain (16 passages) (504), or the more highly passaged HCMV strains, including AD169 (54 passages) (239), and Towne (125 passages) (645).
2.5.2 Ultrastructural Morphology of CyCMV

CyCMV-infected MRC-5 cells were harvested and analyzed by transmission electron microscopy in an effort to visualize the ultrastructural morphology of the newly isolated virus. Classical CMV nucleocapsids (NC) were observed in CyCMV-infected cells and were similar in size (~100-115 nm in diameter) to HCMV and CyCMV Ottawa strain nucleocapsids (22, 294) (Figure 2-1A). Like HCMV and CyCMV Ottawa strain, the structure of the CyCMV Mauritius strain nucleocapsids was a lightly stained viral capsid, which appeared to contain viral DNA, as visualized by densely stained centers. At a lower magnification, viral factories containing nucleocapsids were found in the nucleus of the infected cells, and were closely associated with the tubular structures of the filaments (Figure 2-1B). As the viral infection proceeded to the later phase of infection, the cells began to die, became fragmented, and the nuclear membrane was ruptured. This can be observed in Figure 2-1B in which the nuclear membrane was disrupted in the lower portion of the micrograph.

![Figure 2-1 Electron Micrograph of CyCMV-Infected MRC-5 Cells.](image)

A) CyCMV nucleocapsids (NC) with densely stained DNA-containing centers were visible at 50,000X direct magnification. B) Viral factories containing nucleocapsids were present in the nucleus of infected cells at 40,000X direct magnification. NC: nuclear capsids, NM: nuclear membrane, FL: filamentous structures, nm: nanometer.
2.5.3 Gross Viral Genome Analysis

To assess the purity of the isolated viral DNA and to examine gross viral genome structure, CyCMV was digested with BamHI and HindIII restriction enzymes and fragmented on an agarose gel. The viral DNA was digested into distinct bands with visible differences between cut sites for each restriction enzyme. Furthermore, the isolation resulted in a clean viral preparation with limited cellular genomic DNA contamination. In comparison to BamHI or HindIII digested CyCMV Ottawa (reviewed in Chapter 3; Figure 3-2) (504) and RhCMV 68.1 genomes (337), CyCMV Mauritius strain exhibits a similar band pattern and organization, with a number of separated large fragments (>10kb) observed when digested with HindIII. Once the complete CyCMV genome has been characterized, the sequence will enable prediction of restriction digest cut sites, which will serve to confirm the corresponding digested viral DNA fragment sizes, and allow for the identification of possible discrepancies between the sequenced genome and the isolated viral DNA.
Figure 2-2 Restriction Enzyme Digestion of CyCMV Genome. CyCMV viral DNA was digested with BamHI and HindIII restriction enzymes, and the fragmented DNA (900ng) was separated by electrophoresis on 0.8% agarose gel. Lane MW, 1 kb ladder.

2.5.4 Phylogenetic Analysis of Conserved Genes

To further characterize CyCMV, two highly conserved CMV genes, DNA polymerase and glycoprotein B, were PCR amplified and sequenced in an effort to assess the phylogenetic relationship between CyCMV Mauritius strain and other previously characterized CMVs (Figure 2-3). As expected, DNA polymerase (UL54) was highly conserved between both strains of CyCMV (Ottawa and Mauritius), which were grouped with RhCMV strains, and the most distant relationship was with CMVs isolated from humans and chimpanzees (Figure 2-3A). Similarly,
the phylogenetic analysis of glycoprotein B (gB; UL55) resulted in CyCMV (Mauritius strain) being grouped with CMVs from Old World Monkeys (OWMs) however, within the OWMs CyCMV did not cluster as predicted. Based on the number of substitutions per site, the gB protein sequence of CyCMV Mauritius strain was more closely related to both RhCMV strains than to CyCMV Ottawa strain (Figure 2-3B).

Figure 2-3 Phylogenetic Analysis of DNA Polymerase and Glycoprotein B. Unrooted phylogenetic trees were created using Geneious Tree Builder with the protein sequences for DNA polymerase (A) and glycoprotein B (B). The relationships between various CMV strains is shown as cladograms, and the number of substitutions per site is listed on each branch. The CMV strains include CCMV (chimpanzee CMV), HCMV (AD169), BaCMV (baboon CMV), CgueCMV (colobus guereza CMV; 1.1 and 1.2), CyCMV (Ottawa and Mauritius), and RhCMV (68.1 and 180.92).

Based on the observations from the phylogenetic analysis, a closer examination of the gB protein sequence was performed (Figure 2-4). An inter-strain alignment of gB from both CyCMV strains demonstrated 77.1% identity, and an intra-strain comparison between gB sequences of CyCMV isolated from animals of Filipino origin (Ottawa strain; N=3) versus Mauritius origin (Mauritius strain; N=4) revealed highly conserved intrastrain amino acid sequences with 99.6% and 99.9% identity, respectively (Figure 2-4). In comparison to HCMV, both Ottawa (58%) and Mauritius (59%) strain share similar identity with HCMV gB (22, 504). The gB protein sequence of CyCMV Ottawa strain was less homologous to that of RhCMV 68.1
(77.7% identity) (22) and RhCMV 180.92 (76.6% identity), which is in contrast to the higher degree of sequence homology between CyCMV Mauritius strain and RhCMV 68.1 (88.4% identity), or RhCMV 180.92 (87.9% identity). The differences appear to be localized to the 5’ end of the amino acid sequence, which corresponds to the extracellular domain of gB and the more conserved sequences were at the 3’ end of the protein that functions as the transmembrane domain (Figure 2-4) (22). These results suggest that these observed differences between Mauritius and Ottawa strain represent evolutionary changes in the gB protein.

2.5.5 Growth Kinetics of CyCMV

To examine viral replication in various fibroblast cell lines, a two-step growth kinetics assay was performed. Primary isolates of CyCMV were initially propagated in MRC-5 cells, however as the virus became more adapted to tissue culture the virus was also propagated in an immortalized cynomolagus macaque fibroblast cell line (MSF-T) that was generated in the MacDonald laboratory (23), and in a telomerase-immortalized rhesus macaque fibroblast (Telo-RF) cell line generously provided by Drs. Peter Barry and William Chang (University of California, Davis) (432). In order to evaluate the growth of CyCMV in MSF-T, Telo-RF and MRC-5 cells, the cells were infected with CyCMV at two different MOIs (0.3 and 0.03). The cells and supernatants were collected at days 2, 4, 6, 8, 10 and 12 post-infection and viral titers were determined by standard plaque assay using MRC-5 cells (Figure 2-5). The virus titer in MSF-T cells infected at 0.3 MOI reached its peak by 6 days post-infection (dpi). The peak titre of the CyCMV-infected MSF-T cells was approximately tenfold lower than the peak viral titre observed in Telo-RF (4 dpi) and MRC-5 (12 dpi) cells, which were comparable to each other (Figure 2-5a). The MSF-T cells infected at an MOI of 0.03 reached a maximum viral titer by 8 dpi, and the Telo-RF cells showed peak titers at 6 dpi (Figure 2-5b). CyCMV titers continued to increase beyond 12 dpi in MRC-5 cells infected with CyCMV at an MOI of 0.03. The peak titers in MRC-5 cells infected with CyCMV correlated well with the amount of virus inoculation.
Figure 2-4 Comparison of CyCMV Strain Differences in Glycoprotein B Sequences. Amino acid sequences of glycoprotein B were compared between different animals of CyCMV Mauritius strains and Ottawa strains, as well as RhCMV 68.1 and 180.92 strains. Protein sequences were aligned by ClustalW alignment and shaded using Geneious Pro 5.6.6 (229). The shading represents amino acid similarity at a given residue with 100% in black, 80-100% in dark grey, 60-80% light grey, and <60% no shading. GenBank Accession Numbers: CyCMV Mauritius gB (N/A), CyCMV Ottawa gB (HQ198248), RhCMV 68.1 gB (AAB70024), RhCMV 180.92 gB (AAZ80591).
Figure 2-5 Growth Kinetics of CyCMV in Fibroblast Cell Lines. Replication kinetics of CyCMV in MSF-T, Telo-RF and MRC-5 fibroblast cells. The cells grown in 6-well tissue culture plates were inoculated (in duplicates) with CyCMV at an MOI of 0.3 (a) or 0.03 (b). The cells and the culture supernatant were harvested on days 2, 4, 6, 8, 10 and 12 post-infection, and virus titer in each sample were determined by standard plaque assay on MRC-5 cells. The titers on day ‘0’ in each graph represent the titers of the inoculum. Each data point represents the mean titers from two independent cultures, and the error bars indicate the standard deviation. One of the replicates of MSF-T cells infected at an MOI of 0.03 was spilled during harvest on day 12, and the related data are therefore excluded (23).

2.5.6 MHC Class I Downregulation

Cytomegaloviruses characteristically encode a number of immunoevasin genes that serve to evade the immune response and maintain viral persistence in the host species. More specifically, CMVs encode a series of genes (HCMV homologues US2-US11) that function to downregulate major histocompatibility complex class I (MHC-I), which impedes viral clearance by the CD8⁺ T cell response. We have previously demonstrated that CyCMV downregulates surface expression of MHC-I on CyCMV-infected MRC5 cells (22), however to further characterize CyCMV’s ability to downregulate MHC-I at the protein level, we examined MHC-I heavy chain (HC) and β₂-microglobulin (β₂m) protein expression in CyCMV-infected fibroblast cell lines from different species by Western blot. A clear reduction in MHC-I HC and β₂m levels was observed in all CyCMV-infected cells compared to mock-infected cells, with no alteration in the β-actin levels following infection (Figure 2-6). The downregulation of MHC-I protein expression was most pronounced in Telo-RF-infected cells, which likely reflects viral adaptation
to this cell line that was used for multiple passages during propagation. Genomic sequencing of the CyCMV genome will provide further insight into which immunoevasin genes CyCMV encodes and allow for functional characterization of CyCMV-specific immunomodulatory genes.

![Figure 2-6 MHC Class I Downregulation on CyCMV-infected Cells](image)

**Figure 2-6 MHC Class I Downregulation on CyCMV-infected Cells.** MRC-5, Telo-RF and MSF-T cells. The mock- or CyCMV-infected cells were harvested at 5 days post-infection and analyzed by Western blot for total MHC class I heavy chain (HC), $\beta_2$-microglobulin ($\beta_2$m) and $\beta$-actin expression. m, mock-infected; i, CyCMV-infected. A representative of three independent experiments is shown (23).

### 2.6 Discussion

The viral isolation and characterization of cytomegalovirus isolated from cynomolgus macaques represents a novel non-human primate (NHP) model to study this ubiquitous pathogen. It has been observed that greater than 95% of captive-bred NHPs are CMV-seropositive (26), which is comparable to our evaluation of the CyCMV-seroprevalence in the cynomolgus macaque colony at the Public Health Agency of Canada (PHAC) demonstrating close to 100% seropositivity (22). The cynomolgus macaque colony at PHAC was comprised of animals originating from Indonesia/Philippines, as well as animals imported from the Mauritius Islands. In this regard, CyCMV Ottawa strain was isolated from cynomolgus macaques of Indonesian/Filipino origin (22), and CyCMV Mauritius strain was derived from Mauritius cynomolgus macaques. These two strains have been characterized in parallel allowing for a comparison of the similarities and difference between CyCMV isolated from macaques of different geographic origin.
The relationship between Indonesian versus Mauritian cynomolgus macaques is unique in that Mauritius cynomolgus macaques represent a small founder population that diverged from Indonesian cynomolgus macaques approximately 500 years ago (456, 806, 807, 830). Following their segregation, the population of Mauritian cynomolgus macaques expanded to establish a relatively genetically homogenous population on the island (408). In this regard, evolutionary changes in CyCMV may have occurred, as these viruses characteristically evolve with their host over time. The two strains of CyCMV are comparable in a number of different ways, however they differ with respect to sequence homology in the glycoprotein B coding region of the viral genome (Figure 2-4). The gB amino acid sequence of CyCMV Mauritius strain was more closely related to RhCMV (strains 68.1 and 180.92) than to CyCMV Ottawa strain, suggesting that the virus has evolved from the initial cynomolgus macaque founder population. Glycoprotein B is a highly immunogenic surface protein and a primary target of the neutralizing antibody response (435), which is typically highly conserved. However in the case of the Ottawa strain, the observed gB sequence divergence may be the result of selective pressure from the host immune response, as the majority of the sequence variances fall within the extracellular domain of the protein (Figure 2-4). This hypothesis is supported by a previous study that observed that the majority of the immunodominant epitopes of RhCMV are located within the extracellular region of gB (444, 905). The immunodominant target regions of HCMV have been more thoroughly characterized, with the antibody-binding sites for neutralizing and non-neutralizing antibodies referred to as antigenic domains (AD-1 through AD-5) (542, 780, 851). However, our comparison of the HCMV described antigenic domains revealed little similarity between the HCMV-defined regions and CyCMV or RhCMV. Thus, a more detailed analysis of the immunodominant regions and antibody-binding sites encoded in CyCMV gB will be required before we can conclude that the sequence variation between the two strains of CyCMV can be attributed to immune escape. Furthermore, the complete genomic sequencing will aid in determining if the evolutionary differences are unique to this open reading frame or if they are represented throughout the genome.

CyCMV urine co-cultures were initially grown on human MRC-5 cells, as the doubling rate of the primary cell line was able to support the slow viral replication kinetics of the primary CyCMV isolates. However, as the virus became progressively more lab-adapted, the growth kinetics increased and CyCMV was able to productively replicate in rhesus macaque-derived
immortalized fibroblast cells (Telo-RF). Previous reports have documented that CMV replication was more efficient and had enhanced sustainability in host-derived immortalized fibroblast cell lines compared to primary cells (129). As such, we established an immortalized cynomolgus macaque fibroblast cell line (MSF-T) by transducing primary dermal fibroblasts isolated from a 13-year-old male cynomolgus macaque, with a retrovirus vector expressing human telomerase reverse transcriptase (hTERT) (23). The development of a cynomolgus macaque-specific fibroblast cell line will minimize the potential that CyCMV fibroblast tropism genes and host-specific genes will be deleted from the viral genome during in vitro propagation. In addition to its application for CyCMV propagation, this novel cell line will be of great value to investigators using cynomolgus macaques as a model for other biomedical researcher areas, such as infectious disease, or transplant studies. Indeed, we have already had requests for this cell line from infectious disease researchers working with cynomolgus macaques, and we are currently in the process of making this cell line readily available through the American Tissue Culture Collection (ATCC) organization.

Next-generation sequencing and genetic characterization of the viral genome will allow for an extensive analysis of CMV gene families, and a thorough comparison between CyCMV, RhCMV, and HCMV. Ultimately, genomic characterization will facilitate the development of CyCMV as an HIV vaccine vector through the identification of possible target locations for the insertion of the vaccine antigens and furthermore, it may also aid in elucidating the mechanisms of host-restriction that prevented RhCMV infection in cynomolgus macaques. Next-generation sequencing has been performed in parallel for both CyCMV strains, however the Ottawa strain exhibited the lowest degree of complexity for bioinformatic assembly, thus I proceeded initially with the genomic characterization of the Ottawa strain (Chapter 3). Efforts are currently underway by others students in the MacDonald laboratory to complete the open reading frame assignment and genomic characterization of the Mauritius strain of CyCMV.

ACKNOWLEDGMENTS

We thank Battista Calvieri and Steven Doyle at the Microscopy Imaging Facility at the University of Toronto for their technical assistance. A.K.M. was supported by The Queen Elizabeth II/Community Health Graduate Scholarships in Science and Technology, and the Dr. Bernhard Cinader Graduate Scholarship in Immunology. D.O.W. was supported by a Junior
Investigator Development Award from the Ontario HIV Treatment Network (OHTN). A.A. and K.S.M. also received funding from the OHTN as a postdoctoral fellowship and career scientist award, respectively. This research was funded in part by the Canadian Institutes of Health Research (CIHR) and by the Canadian Foundation for AIDS Research (CANFAR). We sincerely thank the veterinary and technical staff at the NHP colony of Health Canada.
Chapter 3

3 Genomic Sequencing and Characterization of Cynomolgus Macaque Cytomegalovirus

Angie K. Marsh1,3, David O. Willer2,3, Aruna P. N. Ambagala2,3, Misko Dzamba4, Jacqueline K. Chan3, Richard Pilon5, Jocelyn Fournier6, Paul Sandstrom5, Michael Brudno4, Kelly S. MacDonald1,2,3

1Department of Immunology, University of Toronto, ON, Canada, M5S 1A8 2Department of Medicine, University of Toronto, ON, Canada, M5S 1A8 3Department of Microbiology, Mount Sinai Hospital, ON, Canada, M5G 1X5 4Department of Computer Science, University of Toronto, ON, Canada, M5S 1A8 5National HIV & Retroviology Laboratories, National Microbiology Laboratory, Public Health Agency of Canada, Ottawa, ON, Canada, K1A 0L2 6Scientific Services Division, Health Products & Food Branch, Health Canada, Ottawa, ON, Canada K1A 0L2

Author Contributions: All experiments and figures described in this chapter were performed and generated by myself (A.K.M.), with the exception of the bioinformatics, which was performed by M.D., an undergraduate project student in M.B.’s laboratory. M.D. did the bioinformatic assembly, gap closing, and error correction (as described in the methods sections 3.4.3, 3.4.4, 3.4.6). J.K.C. is the technician in the K.S.M. laboratory who helped to prepare and maintain reagents, resources, etc. A.K.M., A.P.A., D.O.W. and K.S.M. were involved with conceiving, designing, and analyzing the experiments. R.P., J.F., P.S. provided the cynomolgus macaque urine samples from our collaborative studies being conducted at the non-human primate colony in Ottawa. Furthermore, the next-generation sequencing was performed at The Centre of Applied Genomics, Toronto, ON (as described in section 3.4.2).

3.1 Chapter Overview

Having isolated and performed preliminary genomic classification of the novel virus as a cynomolgus macaque cytomegalovirus, I proceeded with sequencing and annotating the complete viral genome. I assigned open reading frames for all genes encoded in the genome and performed an extensive analysis of these genes. The detailed genetic characterization of CyCMV will facilitate targeted recombination for vaccine construction, assist with possible attenuation strategies, and also provide information regarding the presence or absence of key genes involved in viral tropism, pathogenesis, and immune evasion, which will be of great importance for downstream in vivo studies.

Main Objective: To isolate, genetically characterize, and develop cynomolgus macaque cytomegalovirus (CyCMV) as an HIV/SIV viral vector for evaluation in a non-human primate model.

Specific Objectives

1. Isolation and characterization of cytomegalovirus from cynomolgus macaques (Chapter 2)
2. Identification and preliminary classification of the isolated virus as a novel CMV (Chapter 2)
3. Next-generation DNA sequencing of CyCMV
4. Bioinformatic assembly of the CyCMV DNA sequence reads
5. Genomic Analysis of CyCMV
6. Cloning of CyCMV as a Bacterial Artificial Chromosome (Chapter 4)
3.2 Abstract

Cytomegalovirus (CMV) is the most common opportunistic infection in immunosuppressed individuals, such as transplant recipients or people living with HIV/AIDS, and congenital CMV is the leading viral cause of developmental disabilities in infants. Due to the highly species-specific nature of CMV, animal models that closely recapitulate human CMV (HCMV) are of growing importance for vaccine development. Here, I present the genomic sequence of a novel non-human primate CMV from cynomolgus macaques (*Macaca fascicularis*; CyCMV). CyCMV (Ottawa strain) was isolated from the urine of a healthy captive-bred four-year old cynomolgus macaque of Filipino origin and the viral genome was sequenced using next-generation Illumina sequencing to an average of 516-fold coverage. The CyCMV genome is 218,041 bp in length, with 49.5% G+C content and 84% protein-coding density. I have identified 262 putative open reading frames (ORFs) with an average coding length of 789 bp. The genomic organization of CyCMV is largely collinear with that of rhesus macaque CMV (RhCMV). Of the 262 CyCMV ORFs, 137 are homologous to HCMV genes, 243 to RhCMV 68.1 and 200 to RhCMV 180.92. CyCMV encodes four ORFs that are not present in RhCMV strains 68.1 or 180.92 but have homologies with HCMV (UL30, UL74A, UL126, UL146). Similar to HCMV, CyCMV does not contain the RhCMV-specific viral homologue of cyclooxygenase-2 (COX-2). This newly characterized cytomegalovirus may provide a novel model in which to study CMV biology and HCMV vaccine development.
3.3 Introduction

Human cytomegalovirus (HCMV), also known as human herpesvirus-5 (HHV-5), is a member of the Betaherpesvirus family (including HHV-6 and HHV-7). Cytomegalovirus (CMV) is a double-stranded DNA virus with the largest genome of all the herpesviruses. The virus is transmitted horizontally through bodily secretions and can cross the placental barrier to facilitate vertical transmission (reviewed in (553)). Cytomegalovirus results in a lifelong infection characterized by the establishment of latency in myeloid progenitor cells followed by periodic reactivation. Cytomegalovirus elicits a strong cellular immune response and in some individuals their CMV-specific T-cells can account for greater than 10% of their total T-cell population (232, 297, 810). In immunocompetent individuals, CMV infection is generally asymptomatic and controlled by the cell-mediated immune response; however in immunocompromised individuals (i.e. neonates, transplant patients and AIDS patients) it can cause severe diseases, such as congenital disorders, CMV retinitis and a variety of opportunistic infections.

Various lab-adapted and clinical strains of HCMV have been isolated and sequenced, most notably are AD169 (131), Toledo (572), Towne (233), Merlin (218). Furthermore, there are a number of clinical strains that have been cloned as bacterial artificial chromosomes (BACs), such as TB40/E (763) and TR, PH, FIX (VR1814) (572). Full-length CMV genomes have been isolated and sequenced from a number of different animal species including mice (671), rats (846), guinea pigs (405, 731) and tree shrews (42). Given their high degree of genetic relatedness to humans, non-human primates (NHPs) likely represent the best animal model to study HCMV biology. A variety of CMVs from Old and New World primates have also been described (457) including chimpanzee CMV (198, 405), rhesus CMV strains 68.1 and 180.92 (337, 691), cercopithecine herpesvirus 5 (CeHV-5) strains GR2715 and Colburn (Accession # FJ483968 and FJ483969, respectively), squirrel monkey CMV (SsciCMV-1; Accession # FJ483967) and owl monkey CMV (AtriCMV-1; Accession # FJ483970). Cytomegaloviruses are highly species-specific viruses (399, 553) and consequently incapable of infecting even closely related species (Marsh, Ambagala et al. submitted; reviewed in Section 1.5.1 Species-Specificity of Non-Human Primate Cytomegaloviruses). This specificity restricts the study of CMV to its target species and reiterates the importance of developing animal models that are closely related to humans in an effort to study HCMV pathogenesis.
Animal models to study CMV biology have been largely limited to mice, guinea pigs and rhesus macaques. As an alternative, cynomolgus macaques (*Macaca fascicularis*) are a species of Old World monkeys that have the potential to serve as a novel NHP model to study CMV pathogenesis. Cynomolgus macaques are extensively used as an animal model for infectious disease research (118, 303, 455, 839), transplant research (259, 419, 734), and are becoming an increasingly popular NHP model for HIV vaccine development (50, 596, 833). Within the field of HIV vaccine design there is a real need to diversify the pool of vectors undergoing testing. Recent studies using rhesus macaque cytomegalovirus (RhCMV) as an SIV vaccine vector have shown much promise in the ability of the vaccine to mount a robust effector memory response, thus providing the vaccinated macaques with long-term protection from SIV disease progression (334, 338) (reviewed in Section 5.1.2 Rhesus Macaque Cytomegalovirus-Based Viral Vector). Cytomegalovirus strains are not conserved between even closely related NHPs and our recent experience suggests that cynomolgus macaques are not readily infected with RhCMV (Marsh, Ambagala et al. submitted; reviewed in Section 1.5.1 Species-Specificity of Non-Human Primate Cytomegaloviruses). In order to overcome this strong host restriction and evaluate CMV as an HIV viral vector in a cynomolgus macaque/SIV model, we must use a cynomolgus macaque cytomegalovirus. We have recently isolated and characterized a novel cytomegalovirus derived from a cynomolgus macaque (CyCMV; Ottawa strain) (22). Here I describe the complete genomic sequence and organization of the CyCMV genome for its use as an alternative NHP model to evaluate CMV pathogenesis and vaccine strategies. I compare and contrast the structural and functional genes of CyCMV to HCMV as well as RhCMV with respect to pathogenesis, immune evasion, and species-specificity.

### 3.4 Materials and Methods

#### 3.4.1 CyCMV Viral DNA Isolation

Cynomolgus macaque cytomegalovirus (CyCMV) Ottawa strain was isolated from catheterized urine samples collected from a healthy captive-bred four-year old cynomolgus macaque of Filipino origin, as described previously (22). A cynomolgus macaque fibroblast cell line was not available at the time of isolation. Initial attempts were made to grow CyCMV in telomerase-immortalized rhesus macaque fibroblast cells (Telo-RF) (432) however, given the rapid growth properties of Telo-RF cells, this cell line could not support the slow growth kinetics
of the CyCMV clinical isolate (22). To circumvent this, I propagated the virus in human fetal lung fibroblast cells (MRC-5) (381), which have slower growth properties and have been used extensively to propagate CMVs (74, 237, 809). Cynomolgus macaque cytomegalovirus was passaged 16 times in MRC-5 cells to obtain high titer virus stocks. The virus was not plaque purified and thus, the sequence likely represents a consensus of one or more strain variants. In order to isolate viral DNA, CyCMV-infected cells were lysed with HIRT extraction buffer (2X TE and 1.2% SDS) for 20 min at RT, treated with 1M sodium chloride overnight at 4°C and subsequently centrifuged at 27,000 x g for 35 min at 4°C to precipitate the cellular DNA and proteins. The supernatant containing viral DNA was treated with an RNAse cocktail (60ug/ml RNAse A and 160U/ml RNAse T1; Fermentas) for 2 h at 37°C and Pronase (1mg/ml; Roche) for 2 h at 37°C. The supernatant was deproteinized with three phenol/chloroform extractions, and the viral DNA was precipitated with 0.3M sodium acetate and two volumes of absolute ethanol overnight at -20°C. The sample was centrifuged at 17,000 x g for 30 min at 4°C and the viral DNA was overlaid on a discontinuous 5-20% sucrose gradient containing ethidium bromide (2ug/ml). Following centrifugation at 200,000 x g for 2.5 h at 4°C, the viral DNA was visualized by UV-illumination, collected and diluted in 1.5 volumes of water and precipitated with 0.3M sodium acetate and 2.5 volumes of absolute ethanol overnight at -20°C. The sample was centrifuged at 15,000 x g for 30 min at 4°C, washed once with 70% ethanol and the viral DNA was resuspended in water. CyCMV viral DNA (1ug) was digested with 20U of HindIII or BamHI at 37°C O/N and fractionated by gel electrophoresis on a 0.8% agarose gel.

3.4.2 Next-Generation DNA Sequencing

Using 9.4 ug of CyCMV DNA, a paired-end library with a 500 bp insert size was prepared to generate read-lengths of 72 bps. To sequence the complete CyCMV genome, high-throughput Illumina Genome Analyzer II paired-end sequencing was performed at The Centre of Applied Genomics, Toronto, ON.

3.4.3 Bioinformatic Assembly

The CyCMV genome was assembled *de novo* from 18,205,114 paired 72 bp reads (~6000-fold coverage) derived from a run of the Illumina Genome Analyzer II platform. Isolated paired-ends were filtered to match the barcode (3,391,350 paired reads, ~1120-fold coverage)
and were assembled using Velvet (version 0.7.55) (909). Best results were obtained using a kmer length 39, shortPaired mode, insert length of 500 bp, and expected coverage of 242, to yield a single large contig of 220 Kbp.

3.4.4 Gap Closing

The resulting Velvet assembly had 11 gaps (runs of Ns), with lengths from 9 to 124. We implemented a simple greedy assembly program that started from a seed sequence, identifying all possible overlapping reads and extended the seed until no further extension was possible. This process is analogous to those described previously (510, 866). By providing the Velvet program with the areas close to gaps as seeds, we were able to generate sequence, and close 10 of the 11 gaps in the initial Velvet assembly.

3.4.5 PCR Sequencing

To confirm the integrity of the sequence, areas of low coverage from the next-generation sequencing data were verified by Sanger sequencing. PCR amplicons were gel-purified with GENECLEAN II (MP Biomedicals) if necessary, cloned into the pCR-Blunt II-TOPO vector (Invitrogen) and transformed into chemically competent cells (Invitrogen). The inserts were confirmed by Sanger sequencing using standard sequencing primers (M13 For (UP): 5’-CACGACGTGTAAAAACGAC-3’ and M13 Rev (-27) - Invitrogen: 5’-CAGGAAACAGCTATGAC-3’). Any remaining sequence was determined by primer-walking. Sequences were aligned by ClustalW alignment using Geneious Pro 5.1.7 (228).

3.4.6 Error Correction

In order to identify additional assembly errors we aligned all of the Illumina reads to the finished assembly and corrected 64 positions (out of total assembly size of 220 Kbp) where the base pair present in the reference genome occurred 7 times less frequently than an alternative base, and replaced each such base with the alternative. Furthermore, for regions of the assembly with low coverage or with conflicting base calls (including all 11 of the gaps initially identified above) we generated 71 Sanger sequences. By analyzing the Sanger sequences together with the base qualities in the aligned Illumina data we were able to correct 24 additional assembly errors (14 single bp modifications and 10 insertion/deletions of 1-23 bp). Finally, the CyCMV sequence
was aligned with that of RhCMV 68.1 (Accession # AY186194) to confirm the orientation of the sequence.

### 3.4.7 Open Reading Frame Assignment

Open reading frames (ORFs) were identified on both strands using Geneious Pro (v 5.1.7) (228) with the following criteria: i) the ORF began with a start codon (ATG) and ended with a stop codon, ii) minimum of 100 aa long (including the start and stop codons), iii) did not overlap another ORF on the same strand within the same reading frame (131). The majority (83%) of the genes were annotated with the above criteria and the remaining genes were identified using a smaller size criterion of ≥30 aa in an effort to identify all possible ORFs. Homology for each ORF was determined using the BLASTP program (NCBI) with >20% identity. BLAST scores (21) and homology details were determined from NCBI.

### 3.4.8 Accession Number

The fully annotated and complete nucleotide sequence of CyCMV (Ottawa strain) has been submitted to GenBank and assigned the accession number: JN227533. The genome homology of CyCMV with other CMVs was determined by Geneious alignment using a global alignment with free and end gaps and a cost matrix of 65% similarity (5.0/-4.0), a gap open penalty of 12 and a gap extension penalty of 3 (228).

### 3.5 Results

#### 3.5.1 Genomic Analysis of CyCMV

Cynomolgus macaque cytomegalovirus (CyCMV) was isolated from a cynomolgus macaque of Filipino origin and sequenced using paired-end sequencing (Illumina) to an average of 516-fold coverage/nucleotide. The CyCMV genome is 218,041 bp in length, with 49.5% G+C content and 84% protein-coding density. CyCMV is shorter and less GC-rich than HCMV AD169 (229,354 bp; 57.2%) (131), chimpanzee CMV (CCMV) (241,087 bp; 61.7%) (198), and more comparable to both strains of RhCMV (68.1: 221,459 bp and 180.92: 215,678 bp; 49% G+C content) (337, 691) (Figure 3-1). Similar to other CMV genomes, CyCMV has a low G+C content at the beginning of the genome (4,559 bp - 17,305 bp), at various regions across the genome (72,074 bp - 74,174 bp; 91,077 bp - 94,827 bp; 137,006 bp - 140,096 bp; 161,230 bp -
163,989 bp; 172,919 bp - 175,284 bp) and at the end of the genome (205,859 bp - 211,243 bp). Consistent with other cytomegalovirus genomes, CyCMV is organized with a unique long region followed by a unique short region (reviewed in Figure 1-12). When comparing the genome sequence and gene products of CyCMV to other cytomegaloviruses, CyCMV most closely resembles RhCMV. The CyCMV genome is 54.8% identical to HCMV AD169 (131), 53.6% CCMV (198), 54.2% AtriCMV-1 (Accession #FJ483970), 57.5% SsciCMV-1 (Accession #FJ483967), 67.5% CeHV-5 strain GR2715 (Accession #FJ483968), 67.8% CeHV-5 strain Colburn (Accession #FJ483969), 89.8% RhCMV 68.1 (337) and 88.2% RhCMV 180.92 (691) at the nucleotide level. The first base call of CyCMV corresponds to base pair 3,996 of HCMV (AD169), -490bp from the first nucleotide of CCMV, and -50bp from beginning of the RhCMV genomes.

Figure 3-1: G+C Content of the CyCMV Genome. The base composition across the CyCMV genome is represented by percent GC (blue) or AT (green) content.

3.5.2 Restriction Enzyme Digestion

To assess gross viral genome structure, a restriction enzyme digest analysis was performed and digested bands were confirmed based on predicted fragment sizes (228). The CyCMV genome was digested with restriction enzymes HindIII and BamHI and fractionated on an agarose gel (Figure 3-2a). The digested CyCMV fragments were compared with the predicted fragments generated from the sequence data using bioinformatics software (Figure 3-2b). All fragments were present at the expected size with the exception of an additional band running at approximately 2.7 Kbp when digested with BamHI (Figure 3-2a). It is possible that there is an extra BamHI site located between cy92 and cyUL69 that is not accurately represented in the final sequence data. This region encompasses the putative origin of lytic replication, a region known for structural complexity (88). This region is inherently challenging to sequence due the presence of inverted and repeated sequence motifs (88) and proved difficult to sequence in our hands by
both next-generation and Sanger sequencing. The RhCMV genome does indeed code for a BamHI restriction site that cuts in the RhCMV origin of lytic replication. I predict that the additional CyCMV band is the result of a missing restriction site in the sequence, although the restriction site may be present in the viral DNA.

**Figure 3-2 Restriction Enzyme Digestion of CyCMV Genome.** To assess gross viral genome structure, a restriction digest analysis was performed. CyCMV viral DNA was digested with HindIII and BamHI restriction enzymes. DNA fragments (900ng) were separated by electrophoresis on 0.8% agarose gel (a). The CyCMV digested DNA has an additional BamHI band (*) at approximately 2.7 Kbp. A map of the CyCMV genome digested with HindIII (31 sites) and BamHI (49 sites) was generated using CLC Main Workbench (v 6.1) (b). Lane MW, 1 kb ladder.

3.5.3 Gene Assignment

I have identified 262 putative open reading frames (ORFs) (Table 3-1), with a mean coding length of 789 bp. The genes were numbered starting at the left of the genome and continuing to the right with similar nomenclature to that used in annotating the ORFs in other non-human primate cytomegaloviruses. The genomic organization of CyCMV is largely collinear with that of RhCMV. Cynomolgus macaque cytomegalovirus gene arrangement with
colour-coded herpesvirus core genes and gene families are shown in Figure 3-3. Herpesviruses share 40 conserved genes, known as core genes, amongst all alpha-, beta-, and gamma-herpesviruses (553). Cynomolgus macaque cytomegalovirus encodes 39 of the 40 core genes with no homologue to HCMV UL108. The function of UL108 is not known and when deleted from the HCMV genome results in only moderate growth defects (233).

Of the 262 CyCMV genes, 137 are homologous to HCMV genes, 243 to RhCMV 68.1 and 200 to RhCMV 180.92. With respect to the RhCMV genomes, CyCMV encodes homologues for 230 of the 260 (89%) RhCMV 68.1 genes and 180 of the 258 (70%) RhCMV 180.92 genes. Cynomolgus macaque cytomegalovirus gene homologues were compared between HCMV and both strains of RhCMV based on their alignment bit score to determine if a particular CyCMV gene is more closely related to a particular strain of CMV (Figure 3-4). The majority of the CyCMV genes are biased towards the RhCMV genomes however, there are some exceptions in which the HCMV gene homologues have the higher bit score. The outliers are mainly membrane proteins (MP) and tegument proteins (TP).
Figure 3-3: Map of Open Reading Frames in CyCMV Genome. CyCMV encodes 262 putative open reading frames (ORFs) that are annotated by gene name and colour coded based on gene families. The genomic organization of CyCMV is largely collinear with that of RhCMV. The CyCMV genes with an HCMV homologue are annotated by “cy” followed by the HCMV name, and the arrowheads indicate the direction of the ORFs. Core genes are herpesvirus core genes.
Figure 3-4: CyCMV Gene Similarities Between RhCMV Strains and HCMV. A bit score was calculated for each CyCMV gene versus its putative homologue in RhCMV or HCMV. The data points would be expected to distribute along the line x = y (gray dashed line), where CyCMV is no more closely related to HCMV than to RhCMV 68.1 or RhCMV 180.92, respectively. The graphs represent comparisons of the CyCMV gene homologue bit scores versus the two RhCMV strains (a), RhCMV 68.1 versus HCMV (b), and RhCMV 180.92 versus HCMV (c). Outlier genes are annotated according to their CyCMV name and putative function. TP = tegument protein, MP = membrane protein, MTMP = multiple transmembrane protein, GPCR = G-protein coupled receptor, MGP = membrane glycoprotein.

3.5.4 Genes Missing From Other Macaque CMVs

Of the 262 CyCMV ORFs, four genes share homology to HCMV genes that are not present in either strain of the RhCMV sequenced genomes (68.1 or 180.92). These genes include UL30, UL74A, UL126 and UL146. It should be noted that a wild-type isolate of RhCMV (RhCMV_{CNPRC}) does encode the HCMV homologue of UL146 (607) and cyUL146 shares 76.1% identity with the RhCMV_{CNPRC} counterpart. The functions of UL30 (cyUL30) and UL126 (cyUL126) have yet to be elucidated, however it is known that UL74A (cyUL74A) is an envelope glycoprotein and UL146 (cyUL146) encodes an alpha-chemokine homologue, vCXCL1, belonging to the UL146 gene family (628). UL146 exhibits a high degree of sequence variability between HCMV strains and amongst species (30). Notably, cyUL146 has retained the chemokine motif, ELRCXC (not shown) that is required for alpha-chemokines to recruit neutrophils (30). During CMV infection, vCXCL1 plays a role in neutrophil attraction and degranulation, resulting in increased viral dissemination both within and between hosts (628).
3.5.5 Tropism Genes

Cynomolagus macaque cytomegalovirus encodes a number of HCMV homologues for tropism genes that have been shown to be essential for HCMV propagation in various cell types. The HCMV homologues (UL128, UL130 and UL131A) that have been shown to be associated with endothelial cell, macrophage and dendritic cell tropism (332, 762) have been retained in CyCMV. The CyCMV genes that share homology with the HCMV UL128-131 region include cyUL128 ex1 (37.1%) and ex2 (55.2%), cyUL130 (40.7%) and cyUL131A (35.3%). An additional HCMV gene known to be required for viral replication in human microvascular endothelial cells (HMVEC) is UL24 (233), for which CyCMV encodes cyUL24 as a homologue with 53.3% identity. Similarly, HCMV UL64 and US29 were shown to be required for growth in human retinal pigment epithelial (RPE) cells (233). Cynomolagus macaque cytomegalovirus encodes a US29 homologue (cyUS29) with 38.1% identity, however does not encode a UL64 homologue. Functional studies will be required to determine if CyCMV can replicate in epithelial cells in the absence of a UL64 homologue. Furthermore, in the functional profiling of HCMV it was determined that the deletion of UL10 and UL16 increases the viral titer in RPE cells, and the deletion of US16 and US19 also results in a higher viral titers in HMVECs (233). Cynomolagus macaque cytomegalovirus does not encode the above listed HCMV homologues with the exception of cyUS19, which is an HCMV homologue of US19. With respect to RhCMV-specific tropism genes, four genes (Rh01, Rh159, Rh160, Rh203) have been shown to be tropism determinants for RhCMV (strain 68.1) replication in rhesus RPE cells (469). The CyCMV homologues for these RhCMV tropism genes include cyTRL1 (84.1%), cyUL148 (89%), cyUL132 (94.9%) and cyUS22 (96.9%). Cynomolagus macaque cytomegalovirus deletion studies will be required to determine if CyCMV exhibits the same impaired viral replication in epithelial cells. The CyCMV tropism genes (cyUL24, cyUL131A, cyUL148, cyUS22) encode full-length homologues for their respective HCMV and/or RhCMV genes. However, the homologues for cyTRL1, cyUL128 ex1/ex2, cyUL130, cyUL132, cyUS19 and cyUS29 only represent partial alignments to the intact CyCMV ORF due to N- and/or C-terminal truncations.

3.5.6 Functional genes

Although CyCMV shares homology with a number of HCMV membrane and tegument proteins, the genes that have a functional role in DNA replication, packaging and egress are the most conserved. These proteins include: DNA packaging terminase component (cyUL89 ex1:
84.8% and ex2: 86.5%, cyUL56: 73.2%), DNA packaging protein (cyUL51: 83.1%), nuclear egress membrane and lamina proteins (cyUL50: 80.6% and cyUL53: 75.4%, respectively), major capsid protein (cyUL86: 76%), single-stranded DNA-binding protein (cyUL57: 74.8%), capsid triplex subunit 1 and 2 (cyUL46: 72.1% and cyUL85: 74%, respectively), DNA polymerase processivity factor (cyUL44: 69.2%), uracil-DNA glycosylase (cyUL114: 69%), DNA helicase primase subunit (cyUL70: 65.3%, cyUL102: 67%, cyUL105: 71.9%), capsid portal protein (cyUL104: 68.3%), viral serine-threonine protein kinase (cyUL97: 66.5%), deoxyribonuclease (cyUL98: 66%), DNA polymerase (cyUL54: 62.2%), small capsid protein (cyUL48a: 65.3%), portal capping protein/DNA packaging (cyUL77: 64.5%) and ribonucleotide reductase subunit 1 (cyUL45: 61.2%). The percent identity to HCMV genes is described and it should be noted that the RhCMV complements for these same genes are even more conserved with an average of 99% identity to their CyCMV counterparts.

3.5.7 Surface glycoproteins

Cytomegalovirus surface glycoproteins are commonly used for identification and classification purposes and to assess phylogenetic relationships between CMVs (22). Cynomolgus macaque cytomegalovirus encodes HCMV homologues for glycoproteins B (cyUL55), N (cyUL73), O (cyUL74), H (cyUL75), M (cyUL100), L (cyUL115). Glycoprotein N (UL73) is a highly variable HCMV glycoprotein (635) however, its CyCMV homologue (cyUL73) exhibits the highest degree of homology (60.8%) compared to the remaining HCMV glycoprotein homologues. Another highly polymorphic HCMV glycoprotein is glycoprotein O (UL74) (635), which is the least conserved among the glycoprotein homologues in CyCMV with 43.7% identity.

3.5.8 Viral Homologues of Chemokine Receptor and G Protein-Coupled Receptor Proteins

Chemokine receptor (CXCL) and G protein-coupled receptor (GPCR) gene homologues are encoded in CMVs from various species. These receptor homologues are organized in gene clusters and the number of repeated genes in a cluster differs between species and between isolates given that these genes are dispensable for growth in fibroblast cells (11). CyCMV has retained six alpha-chemokine receptor homologues that are clustered together in a 3.98 Kbp coding region encompassing cyUL147-cy184. Likewise, CyCMV encodes a cluster of five
GPCR homologues (cyUS28a, cyUS28b, cyUS28c, cyUS28d, cyUS28e) that encode the HCMV homologue of US28, a GPCR known to bind chemokines (848). CyCMV encodes seven genes (cyUL33 ex2, cyUL78, cyUS28a-cyUS28e) that are homologous to three of the four GPCR family genes (UL33, UL78 and US28). The only GPCR homologue absent in the CyCMV genome is US27, a virion envelope glycoprotein (553).

3.5.9 Immunomodulatory Genes

Cytomegaloviruses encode a number of genes that function to evade the immune response of the infected host. Cynomolgus macaque cytomegalovirus encodes HCMV homologues for major histocompatibility complex class I (MHC-I) downregulation genes (US2, US3, US11), viral interleukin-10 (UL111.5A), tumor necrosis receptor homologue (UL144) and anti-apoptotic genes (UL36, UL37 ex1, UL38). In addition to the HCMV genes, cy203 encodes a homologue for the RhCMV-specific gene (rh178) involved in MHC-I downregulation by interfering with the translation of the heavy chain portion of the MHC-I molecule (655, 685). Although originally thought to be unique to RhCMV (655), it appears that this immunomodulatory gene may in fact be an NHP-specific immunoevasin.

The HCMV immunomodulatory genes β2.7, UL16, UL18, UL142, US6, US8 and US10 are not present in CyCMV. Although these genes are important for evading the immune system, they do not show any effects on viral growth in vitro when deleted (233, 529). According to the criteria (>20% identity) used to assign CyCMV homologues, CyCMV does not encode a homologue for the HCMV US6 gene. The RhCMV gene Rh185 shares a low degree of sequence homology with US6, however it has been shown to be functionally similar to US6 and therefore has been assigned as putative homologues (615). Given that cy216 shares significant homology with Rh185 (95.3%) from the RhCMV 68.1 genome, I propose that this CyCMV gene may also function to downregulate MHC-I and may be considered a US6 homologue. We have previously shown that CyCMV downregulates MHC-I expression on CyCMV-infected cells (22). Future functional studies will need to be performed to determine if CyCMV has equivalent or uncharacterized homologues of the missing immunomodulatory genes to evaluate the effects of these deletions on immunomodulation.
3.5.10 Anti-Apoptotic Genes

The anti-apoptotic genes encoded by HCMV include a 2.7 Kbp virally encoded RNA (β2.7) and the UL36-UL38 genes (518). It is likely that CyCMV does not transcribe a β2.7 gene equivalent given than CyCMV does not encode an HCMV homologue for the predicted ORF (RL4) from which the β2.7 transcript is derived (529). The absence of the β2.7 gene does not affect HCMV growth kinetics in vitro (529). Cynomolgus macaque cytomegalovirus encodes homologues for the UL36-UL38 genes (cyUL36 ex1 and ex2, cyUL37 ex1 and cyUL38, respectively). The function of UL36 is to inhibit caspase-8-induced apoptosis (vICA), cyUL36 ex1 and ex2 share 44% and 58.4% identity with UL36. Similarly, UL37 is a mitochondrial inhibitor of apoptosis (vMIA) and cyUL37 ex1 shares 30.5% identity with this HCMV gene. Human cytomegalovirus UL38 (54.7% identity to cyUL38) is an anti-apoptotic gene that blocks the cellular response pathway induced by stress thus preventing cellular apoptosis, alternatively when deleted from the genome the target cells undergo apoptosis and HCMV exhibits viral replication defects (562, 823). It remains to be determined if these CyCMV genes share the same anti-apoptotic roles as their HCMV homologues.

3.5.11 Latency Genes

Of the known cytomegalovirus latency transcripts (CLTs), CyCMV encodes only an HCMV homologue for UL111.5A (cyUL111.5A). The second exon of cy148 (cyUL111.5A ex2) shares 37% identity with the latency-associated UL111.5A HCMV gene that encodes the viral interleukin-10 (vIL-10) protein. Like HCMV, cy148 encodes a spliced transcript with 3 exons, which is analogous to UL111.5A during productive infection (384) and differs from the RhCMV 180.92 homologue (RhUL111a) which encodes 4 exons (691). Comparable to RhCMV, CyCMV does not encode an HCMV homologue for UL81 and thus does not encode the UL81-82 antisense transcript (LUNA) that is involved in latency (56, 675). Although CyCMV has retained a number of gene homologues from the HCMV ULb’ region, it is missing the HCMV homologue for UL138, a known latency gene that is also absent in the RhCMV genomes. Substitution studies have shown that removing UL138 from HCMV does not affect the in vitro growth kinetics of the virus when propagated in fibroblast cells, although they suggest that it could be cell type specific (308).
3.5.12 Spliced Transcripts

Cynomolgus macaque cytomegalovirus encodes at least eight genes that are the products of spliced mRNA transcripts and these include the commonly spliced CMV genes. The spliced transcripts that have two exons include a virion envelope protein/GPCR family (cy52 ex1, cyUL33 ex2), tegument protein vICA (cyUL36), immediate early glycoprotein vMIA (cyUL37), DNA packaging terminase component (cyUL89), early phosphoprotein (cyUL112), IgG Fc-binding glycoprotein (cyUL119 ex1, cy57 ex2) and a putative CC chemokine (cyUL128). Furthermore, cy148 encodes 3 exons (cy148 ex1, cyUL111.5A ex2 and cy148 ex3) to produce the viral interleukin-10 protein. The cy161 ORF is spliced into 4 exons where cyUL122 ex1 and cyUL123 ex2 produce the immediate early proteins 2 and 1, respectively, and cy161 ex3 and cy161 ex4 encode immediate early proteins.

3.5.13 Missing genes

Similar to HCMV, CyCMV does not contain the RhCMV-specific viral homologue of cyclooxygenase-2 (COX-2) (337). CyCMV lacks an approximately 6.7 Kbp coding region equivalent to that of RhCMV rh9-rh16, which encompasses the COX-2 gene homologue encoded by rh10. In comparison to HCMV, CyCMV is lacking full complements for 84 HCMV genes; although the vast majority of these genes are uncharacterized and do not affect viral growth kinetics when deleted from the HCMV genome (Table 3-2). At the left terminus of the genome, CyCMV is lacking all of the RL genes except TRL1 and RL11 (cyTRL1 and cyRL11, respectively). These genes are generally only present in clinical isolates of CMV as they are dispensable for growth in vitro (233). The RL11 family of genes is not present in mouse or rat CMV (671, 846). Of the absent HCMV genes, the only genes that have been reported to be required for viral replication are UL60 and UL90, both functionally uncharacterized proteins. It appears that CyCMV is lacking the HCMV homologues (UL58-UL68) spanning the origin of lytic replication (oriLyt) that resides between cy92 and cyUL69. These HCMV genes (UL58-UL68) are present in the AD169 strain, however they are not present in the Toledo strain of HCMV or in either of the RhCMV strains. The only RhCMV homologues missing from the oriLyt area are rh94 and rh96, suggesting that CyCMV is not lacking any crucial genes in this region and has a similar gene allocation as HCMV (Toledo) and RhCMV.
Furthermore, CyCMV is also lacking complements of the UL2, UL12, UL65, UL108 and UL129 genes, all of which do not have a known function with the exception of UL2 (putative membrane protein). Only a modest effect on viral replication has been observed when these genes are deleted from the HCMV genome (233). There is ambiguity in the literature regarding the effect of knocking out the membrane protein UL124 (233). The remaining HCMV gene deletions have yet to be examined for their effect on viral growth kinetics.

3.5.14 Phylogenetic Analysis

Phylogenetic trees have been generated to depict the evolutionary relatedness of common CMV genes between species (Figure 3-5). Genes that had the greatest number of sequenced strains were included in this analysis. As expected, these CyCMV-encoded genes uniformly group more closely with RhCMV than the other CMV strains. The cy216 gene was included in this analysis to clarify the discrepancy regarding the HCMV US6 homologue. For this gene, the number of substitutions per site for the branches separating CyCMV and RhCMV from HCMV is relatively high (0.994) in comparison to other genes represented by the phylogenetic trees. Furthermore, of interest is the cy216 gene in which the CyCMV complement is more closely related to RhCMV 68.1 than is RhCMV 180.92.

Figure 5. Phylogenetic analysis of CyCMV genes. Unrooted phylogenetic trees were created using Geneious Tree Builder with the protein sequences of various CMV strains. The relationship between strains is shown by cladogram and substitutions per site are listed on each branch. The CMV strains include: CyCMV (bold), HCMV (human CMV AD169), CCMV (chimpanzee CMV), BaCMV (baboon CMV), GoCMV (gorilla CMV), ColCMV (colobus guereza CMV) and RhCMV (rhesus CMV 68.1 and 180.92). SSDB = single-stranded DNA-binding protein.
Figure 3-5: Phylogenetic Analysis of CyCMV Genes. Unrooted phylogenetic trees were created using Geneious Tree Builder with the protein sequences of various CMV strains. The relationships between strains is shown as cladograms, and the number of substitutions per site is listed on each branch. The CMV strains include CyCMV (bold), HCMV AD169, CCMV, BaCMV (baboon CMV), GoCMV (gorilla CMV), ColCMV (colobus guereza CMV) and RhCMV (68.1 and 180.92). SSDB, single-stranded DNA-binding protein.

3.6 Discussion

The newly characterized cynomolgus macaque cytomegalovirus (Ottawa strain) is 218,041 bp in length, encodes 262 open reading frames (ORFs) and is most closely related to the two published genomes of RhCMV (strains 68.1 and 180.92). Although I have predicted 262 ORFs, I acknowledge that this may not be a complete representation of the CyCMV genome and that there may be additional ORFs encoded in the genome that have yet to be elucidated. The virus was not plaque purified and thus, the sequence likely represents a consensus of one or more strain variants. Our particular Illumina sequencing run had a calculated error frequency of 1% to act as a baseline for substitution frequencies. In this manner, I determined that sequence calls other than the consensus represent, on average, only 1.9%. This determination does not allow us to calculate how many deviances from consensus are contained simultaneously in a single genome sequence, but does suggest that our current sequence likely represents a mixed population with only a diminutive portion of variants. With respect to interhost variability, I have previously examined the amino acid sequence of glycoprotein B (cyUL55) and have observed 99% identity between animals of the same geographic origin (22).

As the Ottawa strain of CyCMV is a multiply tissue culture passaged virus, in vitro passage may have resulted in deletions impairing coding capacity. However, in comparison to the multiple passages required to generate the HCMV strains AD169 (54 passages) (239) and Towne (125 passages) (645), CyCMV (16 passages) would be considered only a moderately passaged isolate. Potential gene deletions could be further investigated by sequencing and characterizing a different isolate of CyCMV, specifically a low-passage strain. The CyCMV genome is unique in that it encodes four HCMV homologues (UL30, UL74A, UL126 and UL146) that are not present in either of the published RhCMV genomes (68.1 or 180.92), although a HCMV homologue of UL146 is present in a wild-type strain of RhCMV (RhCMV_{CNPRC}) (607). There is no putative function for UL30 and UL126, however it is known that UL74A is an envelope glycoprotein and UL146 is an alpha-chemokine homologue.
It has been suggested that vCXCL1 may act as a virulence determinant of CMV disease in individuals with a compromised adaptive immune system (551). Although CyCMV is a multi-passaged derivative, with respect to the chemokine and GPCR gene clusters, CyCMV resembles a minimally passaged virus in that it has retained the clusters of six alpha-chemokine receptor homologues and five GPCR homologues. The wild-type isolates of RhCMV encode six CXCL and five GPCR gene clusters, however the annotated RhCMV strains 68.1 and 180.92 have deleted half or all of the CXCL genes while maintaining all of GPCR genes in the clusters (11). CyCMV does not appear to have deleted these genes in the same way that the RhCMV genomes have.

CyCMV does not encode a viral cyclooxygenase-2 (COX-2) gene that appears to be unique to RhCMV (337). Cellular COX-2 expression is induced upon HCMV infection and has been shown to play an important role in HCMV replication (914). Unlike HCMV infection, RhCMV infection does not induce cellular COX-2 expression in the presence of the viral COX-2 isoform encoded in the RhCMV genome (rh10) (708). Future studies will be required to determine if CyCMV infection induces cellular COX-2 expression in the same way as HCMV.

Given the importance of cynomolgus macaques as a widely utilized animal model for infectious diseases and transplant research, the isolation and characterization of this highly prevalent endogenous virus may have a variety of applications. The seroprevalence of CyCMV in the cynomolgus macaque colony at the Public Health Agency of Canada in Ottawa, Canada is estimated to be 100% when measured by a CyCMV-specific ELISA (22). In other studies, it has been observed that greater than 95% of captive-bred NHPs are CMV-seropositive (26), fortunately it has been shown that CMV-seropositive rhesus macaques can be superinfected with RhCMV (335). This has yet to be examined in cynomolgus macaques, although we believe that like RhCMV and HCMV it would be possible achieve superinfection with CyCMV.

Cytomegaloviruses have evolved with their hosts over millions of years and have encoded CMV-specific genes that are related to each host species. We have preliminary data suggesting that it may be inherently difficult to cross the species-specific barrier and infect cynomolgus macaques with RhCMV (Marsh A.K., Ambagala, A.P.N., et al. submitted; refer to Section 1.5.1). Although the CyCMV genome is nearly 90% identical to RhCMV (at the nucleotide level) and the genes are largely collinear with that of RhCMV, clearly there are
factors influencing the host range specificity of the virus. The mechanism by which a host cell restricts viral replication from a different species has not been well elucidated. However, it is known that this restriction does not occur exclusively during the entry phase of CMV infection as it has been shown that CMV has the capacity to infect a host cell from a distant species (254). One possible mechanism by which the host cell inhibits CMV replication from other species may be mediated by apoptosis, suggesting that the foreign virus cannot overcome the cellular innate immune defense of the host (399). Cytomegaloviruses encode anti-apoptotic genes (β2.7, UL36, UL37 ex1 and UL38) that function to overcome the apoptosis response induced by the host innate immune response following CMV infection (553). The HCMV homologues of the UL36-38 anti-apoptotic genes are encoded in the CyCMV genome (cyUL36 ex1 and ex2, cyUL37 ex1 and cyUL38) and these CyCMV genes share high homology (~96-97% identity) with their RhCMV counterparts (Table 3-1). Although it has been suggested that these anti-apoptotic genes are important for host restriction in vivo, this does not appear to be the situation in vitro where the species-specificity is less restricted. CyCMV productively infects and replicates in human (MRC-5), rhesus (Telo-RF) and cynomolgus (MSF-T) macaque fibroblast cell lines (23). It is known that when CMV strains are grown in fibroblast cell lines, they classically eliminate the tropism genes required for replication in different cell types, most notably endothelial cells (855). Although CyCMV has been propagated in fibroblast cells prior to sequencing, the genes that are required for endothelial cell tropism (UL128, UL130 and UL131A) have been retained in CyCMV (cyUL128 ex1/ex2, cyUL130 and cyUL131A, respectively). Furthermore, I have preliminary evidence demonstrating that CyCMV infects and efficiently replicates in human umbilical vein endothelial cells (HUVECs) (as shown in Chapter 4: Figure 4-4). Endothelial cell tropism plays an important role in natural infection and viral transmission (854), thus CyCMV may have utility for examining viral dissemination and pathogenesis in endothelial cells.

Congenital CMV remains the most common viral cause of birth defects in newborns and yet there is still no vaccine (113). The burden of CMV disease is not only apparent in children but also in adults, specifically those receiving solid organ or bone marrow transplants, and those suffering from an immune compromising disease such as HIV/AIDS. We have reason to be hopeful regarding the ability to make a CMV vaccine given the success of another herpesvirus, Varicella Zoster virus (VZV), in which the licenced vaccine has been highly effective in reducing the mortality associated with varicella infections in the United States (503). We hope
this newly sequenced and characterized CyCMV genome will provide the necessary groundwork for future studies evaluating the utility of cynomolgus macaques as an alternative NHP model in which to study CMV biology, pathogenesis and vaccine design. Furthermore, in an effort to develop CyCMV as a viral vector for evaluation as an HIV vaccine, the genomic characterization will elucidate target regions for the cloning of CyCMV as a Bacterial Artificial Chromosome (BAC) to enable downstream insertion of the vaccine antigens and other viral manipulations of the vaccine construct. In addition, the sequenced genome will be critical in guiding second-generation vaccines and attenuation strategies, which may be required to safely translate this vaccine strategy into humans.

ACKNOWLEDGEMENTS

A.K.M. was supported by The Queen Elizabeth II/Community Health Graduate Scholarships in Science and Technology, and the Dr. Bernhard Cinader Graduate Scholarship in Immunology. D.O.W. was supported by a Junior Investigator Development Award from the Ontario HIV Treatment Network (OHTN). A.A. and K.S.M. also received funding from the OHTN as a postdoctoral fellowship and career scientist award, respectively. This research was funded in part by the Canadian Institutes of Health Research (CIHR). We sincerely thank the veterinary and technical staff at the NHP colony of Health Canada.
## Table 3-1: CyCMV Gene Products

<table>
<thead>
<tr>
<th>ORF&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Translation Start</th>
<th>Strand</th>
<th>Size (aa)</th>
<th>Mol mass (kDa)</th>
<th>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</th>
<th>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</th>
<th>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>cyTRL1</td>
<td>1040</td>
<td>+</td>
<td>579</td>
<td>63.6</td>
<td>RL1 family</td>
<td>TRL1 (37)</td>
<td>Rh01 (84.1) rhRL1 (84.9)</td>
</tr>
<tr>
<td>cy02</td>
<td>1693</td>
<td>+</td>
<td>87</td>
<td>9.9</td>
<td></td>
<td></td>
<td>rh02 rh2</td>
</tr>
<tr>
<td>cy03</td>
<td>2097</td>
<td>+</td>
<td>245</td>
<td>26.2</td>
<td></td>
<td></td>
<td>rh2.1 rh2.1</td>
</tr>
<tr>
<td>cy04</td>
<td>2828</td>
<td>+</td>
<td>166</td>
<td>18.2</td>
<td></td>
<td></td>
<td>rh03 rh3</td>
</tr>
<tr>
<td>cy05</td>
<td>2913</td>
<td>-</td>
<td>233</td>
<td>30.0</td>
<td></td>
<td></td>
<td>rh04 rh4</td>
</tr>
<tr>
<td>cy06</td>
<td>2961</td>
<td>+</td>
<td>101</td>
<td>11.5</td>
<td></td>
<td></td>
<td>rh3.1 rh3.1</td>
</tr>
<tr>
<td>cy07</td>
<td>3298</td>
<td>+</td>
<td>162</td>
<td>18.2</td>
<td></td>
<td></td>
<td>rh3.2 rh3.2</td>
</tr>
<tr>
<td>cyRL11</td>
<td>3738</td>
<td>+</td>
<td>274</td>
<td>30.1</td>
<td>IgG Fc-binding membrane glycoprotein/ RL11 family</td>
<td>RL11 (31.3)</td>
<td>Rh05 rh5</td>
</tr>
<tr>
<td>cy09</td>
<td>4873</td>
<td>+</td>
<td>162</td>
<td>18.6</td>
<td></td>
<td></td>
<td>rh06 rh6</td>
</tr>
<tr>
<td>cy10</td>
<td>5475</td>
<td>+</td>
<td>200</td>
<td>22.7</td>
<td></td>
<td></td>
<td>rh07 rh7</td>
</tr>
<tr>
<td>cy11</td>
<td>6049</td>
<td>+</td>
<td>214</td>
<td>24.3</td>
<td></td>
<td></td>
<td>rh08 rh8</td>
</tr>
<tr>
<td>cy12</td>
<td>8067</td>
<td>+</td>
<td>358</td>
<td>39.9</td>
<td></td>
<td></td>
<td>Rh17 rh17</td>
</tr>
<tr>
<td>cy13</td>
<td>8075</td>
<td>-</td>
<td>121</td>
<td>13.1</td>
<td></td>
<td></td>
<td>rh18 rh18</td>
</tr>
<tr>
<td>cyUL7</td>
<td>9222</td>
<td>+</td>
<td>315</td>
<td>35.6</td>
<td>Putative membrane</td>
<td>UL7</td>
<td>Rh19 rhUL7</td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>cyUL6</td>
<td>10224</td>
<td>10817</td>
<td>+</td>
<td>198</td>
<td>22.1</td>
<td>glycoprotein/ RL11 family</td>
<td>(34.7)</td>
</tr>
<tr>
<td>cyUL6</td>
<td>10844</td>
<td>11527</td>
<td>+</td>
<td>228</td>
<td>26.4</td>
<td>Putative membrane glycoprotein/ RL11 family</td>
<td>(27.9)</td>
</tr>
<tr>
<td>cy17</td>
<td>11527</td>
<td>11772</td>
<td>+</td>
<td>82</td>
<td>10.0</td>
<td>Putative membrane glycoprotein/ RL11 family</td>
<td>(28.5)</td>
</tr>
<tr>
<td>cyUL11</td>
<td>11724</td>
<td>12404</td>
<td>+</td>
<td>227</td>
<td>25.4</td>
<td>Membrane glycoprotein/ RL11 family</td>
<td>(33.3)</td>
</tr>
<tr>
<td>cyUL9b</td>
<td>12455</td>
<td>12817</td>
<td>+</td>
<td>121</td>
<td>13.7</td>
<td>Putative membrane glycoprotein/ RL11 family</td>
<td>(25.3)</td>
</tr>
<tr>
<td>cyUL9c</td>
<td>12896</td>
<td>13570</td>
<td>+</td>
<td>225</td>
<td>24.7</td>
<td>Putative membrane glycoprotein/ RL11 family</td>
<td>(28.5)</td>
</tr>
<tr>
<td>cyUL9d</td>
<td>13590</td>
<td>14408</td>
<td>+</td>
<td>273</td>
<td>31.0</td>
<td>Putative membrane glycoprotein/ RL11 family</td>
<td>(31.9)</td>
</tr>
<tr>
<td>cy22</td>
<td>14556</td>
<td>15164</td>
<td>+</td>
<td>203</td>
<td>23.0</td>
<td>UL6</td>
<td>rh20</td>
</tr>
<tr>
<td>cy23</td>
<td>15166</td>
<td>15789</td>
<td>+</td>
<td>208</td>
<td>23.8</td>
<td>UL9</td>
<td>rh21</td>
</tr>
<tr>
<td>cy24</td>
<td>15865</td>
<td>17235</td>
<td>+</td>
<td>457</td>
<td>49.6</td>
<td>RH22</td>
<td>rh22</td>
</tr>
<tr>
<td>cy25</td>
<td>17311</td>
<td>17634</td>
<td>-</td>
<td>108</td>
<td>11.5</td>
<td>UL9&lt;sup&gt;e&lt;/sup&gt;</td>
<td>rh24</td>
</tr>
<tr>
<td>cyUL13</td>
<td>17351</td>
<td>18658</td>
<td>+</td>
<td>436</td>
<td>50.3</td>
<td>Putative secreted protein</td>
<td>(34.2)</td>
</tr>
<tr>
<td>cy27</td>
<td>17573</td>
<td>17743</td>
<td>-</td>
<td>57</td>
<td>6.3</td>
<td>UL13</td>
<td>rh30</td>
</tr>
<tr>
<td>cy28</td>
<td>17798</td>
<td>18025</td>
<td>-</td>
<td>76</td>
<td>8.7</td>
<td>RL11 family</td>
<td>(93.2)</td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>----------------</td>
<td>------------------</td>
<td>------</td>
<td>--------</td>
<td>-----------</td>
<td>---------------</td>
<td>---------------------------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>cyUL14</td>
<td>18934</td>
<td>19845</td>
<td>+</td>
<td>304</td>
<td>35.3</td>
<td>Putative membrane glycoprotein/ UL14 family</td>
<td>UL14 (32.8)</td>
</tr>
<tr>
<td>cy30</td>
<td>20786</td>
<td>21109</td>
<td>+</td>
<td>108</td>
<td>12.7</td>
<td></td>
<td>R35 (95.4)</td>
</tr>
<tr>
<td>cyUL19</td>
<td>21438</td>
<td>21725</td>
<td>+</td>
<td>96</td>
<td>10.9</td>
<td></td>
<td>UL19 (46.2)</td>
</tr>
<tr>
<td>cyUL20</td>
<td>21834</td>
<td>23189</td>
<td>+</td>
<td>452</td>
<td>51.0</td>
<td>T-cell receptor γ chain homologue</td>
<td>UL20 (36.7)</td>
</tr>
<tr>
<td>cyUL21A</td>
<td>23294</td>
<td>23656</td>
<td>-</td>
<td>121</td>
<td>13.9</td>
<td>CC chemokine binding protein</td>
<td>UL21A (41.6)</td>
</tr>
<tr>
<td>cy34</td>
<td>23932</td>
<td>24195</td>
<td>-</td>
<td>88</td>
<td>9.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cyUL23</td>
<td>24752</td>
<td>25690</td>
<td>-</td>
<td>313</td>
<td>35.9</td>
<td>Tegument protein/ US22 family</td>
<td>UL23 (44.4)</td>
</tr>
<tr>
<td>cy36</td>
<td>25437</td>
<td>25817</td>
<td>+</td>
<td>127</td>
<td>14.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cyUL24</td>
<td>25747</td>
<td>26673</td>
<td>-</td>
<td>309</td>
<td>35.1</td>
<td>Tegument protein/ US22 family</td>
<td>UL24 (53.3)</td>
</tr>
<tr>
<td>cyUL25</td>
<td>26740</td>
<td>28509</td>
<td>+</td>
<td>590</td>
<td>67.4</td>
<td>Tegument phosphoprotein/ UL25 family</td>
<td>UL25 (40.4)</td>
</tr>
<tr>
<td>cyUL26</td>
<td>28573</td>
<td>29328</td>
<td>-</td>
<td>252</td>
<td>28.2</td>
<td>Tegument protein; transcriptional activator of major immediate early promoter/ US22 family</td>
<td>UL26 (46.7)</td>
</tr>
<tr>
<td>cyUL27</td>
<td>29282</td>
<td>31015</td>
<td>-</td>
<td>578</td>
<td>65.7</td>
<td></td>
<td>UL27 (55.8)</td>
</tr>
<tr>
<td>cyUL28</td>
<td>31100</td>
<td>32113</td>
<td>-</td>
<td>338</td>
<td>38.7</td>
<td>US22 family</td>
<td>UL28 (67.1)</td>
</tr>
<tr>
<td>cy42</td>
<td>31514</td>
<td>31786</td>
<td>+</td>
<td>91</td>
<td>10.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------</td>
<td>-----------------</td>
<td>--------</td>
<td>-----------</td>
<td>----------------</td>
<td>----------------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>cy43</td>
<td>32224</td>
<td>32388</td>
<td>+</td>
<td>55</td>
<td>6.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cyUL29</td>
<td>32242</td>
<td>33252</td>
<td>-</td>
<td>337</td>
<td>38.9</td>
<td>US22 family</td>
<td>RH29 (62.1)</td>
</tr>
<tr>
<td>cy45</td>
<td>32385</td>
<td>32789</td>
<td>-</td>
<td>135</td>
<td>15.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cy46</td>
<td>32638</td>
<td>33150</td>
<td>+</td>
<td>171</td>
<td>19.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cy47</td>
<td>33304</td>
<td>33537</td>
<td>+</td>
<td>78</td>
<td>9.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cyUL30</td>
<td>33326</td>
<td>33592</td>
<td>-</td>
<td>89</td>
<td>10.3</td>
<td></td>
<td>UL30 (37.7)</td>
</tr>
<tr>
<td>cy49</td>
<td>33582</td>
<td>34016</td>
<td>-</td>
<td>145</td>
<td>17.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cyUL31</td>
<td>33899</td>
<td>35524</td>
<td>+</td>
<td>542</td>
<td>61.0</td>
<td>dUTPase family</td>
<td>UL31 (58.6)</td>
</tr>
<tr>
<td>cyUL32</td>
<td>35535</td>
<td>37667</td>
<td>-</td>
<td>711</td>
<td>79.5</td>
<td>Major tegument protein</td>
<td>UL32 (54.1)</td>
</tr>
<tr>
<td>cy52 ex1</td>
<td>37621</td>
<td>37773</td>
<td>+</td>
<td>51</td>
<td>6.0</td>
<td></td>
<td>UL33 (94)</td>
</tr>
<tr>
<td>cyUL33 ex2</td>
<td>38035</td>
<td>39024</td>
<td>+</td>
<td>330</td>
<td>37.2</td>
<td>Virion envelope protein/GPCR family</td>
<td>UL33 (56.2)</td>
</tr>
<tr>
<td>cyUL34</td>
<td>39229</td>
<td>40086</td>
<td>+</td>
<td>286</td>
<td>32.9</td>
<td>Represses US3 transcription</td>
<td>UL34 (64.4)</td>
</tr>
<tr>
<td>cy54</td>
<td>39543</td>
<td>39929</td>
<td>+</td>
<td>129</td>
<td>13.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cyUL35</td>
<td>40149</td>
<td>41921</td>
<td>+</td>
<td>591</td>
<td>66.9</td>
<td>Tegument phosphoprotein;</td>
<td>UL35</td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>cyUL36 ex1</td>
<td>42038</td>
<td>-</td>
<td>388</td>
<td>44.7</td>
<td>Interacts with UL82 protein/ UL25 family</td>
<td>(43.9)</td>
<td>(98.5)</td>
</tr>
<tr>
<td>cy57</td>
<td>42308</td>
<td>+</td>
<td>100</td>
<td>11.5</td>
<td>Immediate early tegument protein; inhibitor of caspase-8-induced apoptosis (vICA)/ US22 family</td>
<td>rhUL36 (96.9)</td>
<td>rhUL36 (97.2)</td>
</tr>
<tr>
<td>cyUL36 ex2</td>
<td>43248</td>
<td>-</td>
<td>94</td>
<td>10.5</td>
<td>Immediate early tegument protein; inhibitor of caspase-8-induced apoptosis (vICA)/ US22 family</td>
<td>UL36 (44)</td>
<td>rhUL36 (96.6)</td>
</tr>
<tr>
<td>cyUL37 ex1</td>
<td>43627</td>
<td>-</td>
<td>273</td>
<td>31.2</td>
<td>Immediate early glycoprotein; mitochondrial inhibitor of apoptosis (vMIA)</td>
<td>UL37 (30.5)</td>
<td>rhUL37 (96)</td>
</tr>
<tr>
<td>cy59</td>
<td>44646</td>
<td>+</td>
<td>128</td>
<td>13.5</td>
<td></td>
<td>rh63 (96.9)</td>
<td></td>
</tr>
<tr>
<td>cyUL38</td>
<td>44750</td>
<td>-</td>
<td>294</td>
<td>33.2</td>
<td>Virion envelope glycoprotein</td>
<td>UL38 (54.7)</td>
<td>Rh64 (96.6)</td>
</tr>
<tr>
<td>cy61</td>
<td>44816</td>
<td>+</td>
<td>122</td>
<td>13.8</td>
<td></td>
<td>rh5 (93.4)</td>
<td>rh65 (92.6)</td>
</tr>
<tr>
<td>cy62</td>
<td>45377</td>
<td>+</td>
<td>100</td>
<td>11.3</td>
<td></td>
<td>rh5 (89)</td>
<td>rh65 (87.9)</td>
</tr>
<tr>
<td>cyUL37 ex2</td>
<td>45673</td>
<td>-</td>
<td>97</td>
<td>11.0</td>
<td>Immediate early glycoprotein</td>
<td>UL37 (36.6)</td>
<td>rhUL37 (90.6)</td>
</tr>
<tr>
<td>cy63</td>
<td>46313</td>
<td>-</td>
<td>177</td>
<td>19.0</td>
<td></td>
<td>rh67 (93.2)</td>
<td></td>
</tr>
<tr>
<td>cyUL41A</td>
<td>46921</td>
<td>-</td>
<td>80</td>
<td>9.4</td>
<td>Virion envelope protein</td>
<td>UL41A&lt;sup&gt;f&lt;/sup&gt; (36.7)</td>
<td>rhUL41a (99)</td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>----------------</td>
<td>------------------</td>
<td>------------------</td>
<td>--------</td>
<td>----------</td>
<td>----------------</td>
<td>----------------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>cyUL42</td>
<td>47291</td>
<td>47680</td>
<td>-</td>
<td>130</td>
<td>14.3</td>
<td>Putative membrane protein</td>
<td>UL42 (45.9)</td>
</tr>
<tr>
<td>cyUL43</td>
<td>47664</td>
<td>48665</td>
<td>-</td>
<td>334</td>
<td>38.5</td>
<td>Tegument protein/ US22 family</td>
<td>UL43 (46.3)</td>
</tr>
<tr>
<td>cyUL44</td>
<td>48784</td>
<td>49956</td>
<td>-</td>
<td>391</td>
<td>44.0</td>
<td>DNA polymerase processivity factor</td>
<td>UL44 (69.2)</td>
</tr>
<tr>
<td>cy68</td>
<td>49186</td>
<td>49662</td>
<td>+</td>
<td>159</td>
<td>18.0</td>
<td>Tegument protein; ribonucleotide reductase subunit 1</td>
<td></td>
</tr>
<tr>
<td>cyUL45</td>
<td>50197</td>
<td>52749</td>
<td>-</td>
<td>851</td>
<td>97.0</td>
<td>Tegument protein; ribonucleotide reductase subunit 1</td>
<td>UL45 (61.2)</td>
</tr>
<tr>
<td>cy70</td>
<td>50458</td>
<td>50817</td>
<td>+</td>
<td>120</td>
<td>13.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cy71</td>
<td>52235</td>
<td>52537</td>
<td>+</td>
<td>101</td>
<td>11.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cyUL46</td>
<td>52768</td>
<td>53640</td>
<td>-</td>
<td>291</td>
<td>33.1</td>
<td>Capsid triplex subunit 1</td>
<td>UL46&lt;sup&gt;f&lt;/sup&gt; (72.1)</td>
</tr>
<tr>
<td>cyUL47</td>
<td>53639</td>
<td>56515</td>
<td>+</td>
<td>959</td>
<td>110.6</td>
<td>Tegument protein</td>
<td>UL47 (41.8)</td>
</tr>
<tr>
<td>cyUL48</td>
<td>56536</td>
<td>63069</td>
<td>+</td>
<td>2178</td>
<td>246.9</td>
<td>Large tegument protein</td>
<td>UL48 (41.3)</td>
</tr>
<tr>
<td>cy75</td>
<td>57035</td>
<td>57427</td>
<td>+</td>
<td>131</td>
<td>15.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cyUL48a</td>
<td>63141</td>
<td>63362</td>
<td>-</td>
<td>74</td>
<td>8.4</td>
<td>Small capsid protein</td>
<td>UL48a (65.3)</td>
</tr>
<tr>
<td>cy77</td>
<td>63241</td>
<td>63906</td>
<td>+</td>
<td>222</td>
<td>23.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cyUL49</td>
<td>63355</td>
<td>64824</td>
<td>-</td>
<td>490</td>
<td>56.2</td>
<td></td>
<td>UL49 (71.5)</td>
</tr>
</tbody>
</table>

<sup>a</sup> cy: cytomegalovirus; Rh: rhinovirus

<sup>b</sup> US22: unique sequence 22; UL: unique length

<sup>c</sup> HCMV: human cytomegalovirus

<sup>d</sup> Homologue: percentage identity

<sup>e</sup> rh: rhinovirus

<sup>f</sup> UL46: UL46 family
<table>
<thead>
<tr>
<th>ORF</th>
<th>Translation Start</th>
<th>Translation Stop</th>
<th>Strand</th>
<th>Size (aa)</th>
<th>Mol mass (kDa)</th>
<th>Function/gene family</th>
<th>HCMV Homologue (%)</th>
<th>RhCMV Homologue (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cyUL50</td>
<td>64814</td>
<td>65695</td>
<td>-</td>
<td>294</td>
<td>32.4</td>
<td>Inner nuclear membrane protein; nuclear egress membrane protein</td>
<td>UL50 (80.6)</td>
<td>rhUL50 (99.3)</td>
</tr>
<tr>
<td>cyUL51</td>
<td>65721</td>
<td>66056</td>
<td>-</td>
<td>112</td>
<td>12.4</td>
<td>DNA packaging protein; terminase component</td>
<td>UL51 (83.1)</td>
<td>Rh82 (99.1) rhUL51 (97.3)</td>
</tr>
<tr>
<td>cyUL52</td>
<td>66116</td>
<td>67774</td>
<td>+</td>
<td>553</td>
<td>62.5</td>
<td>DNA packaging protein; terminase component</td>
<td>UL52 (55.9)</td>
<td>Rh83 (98.2) rhUL52 (98.4)</td>
</tr>
<tr>
<td>cy82</td>
<td>66226</td>
<td>66534</td>
<td>-</td>
<td>103</td>
<td>10.8</td>
<td>DNA packaging protein; terminase component</td>
<td>UL53 (75.4)</td>
<td>Rh85 (98.3)</td>
</tr>
<tr>
<td>cyUL53</td>
<td>67767</td>
<td>68630</td>
<td>+</td>
<td>288</td>
<td>32.9</td>
<td>Tegument protein; nuclear matrix protein; nuclear egress lamina protein</td>
<td>UL53 (75.4)</td>
<td>rh86 (100)</td>
</tr>
<tr>
<td>cy84</td>
<td>68175</td>
<td>68606</td>
<td>-</td>
<td>144</td>
<td>16.0</td>
<td>DNA polymerase</td>
<td>UL54 (62.2)</td>
<td>Rh87 (99) rhUL54 (98.9)</td>
</tr>
<tr>
<td>cyUL54</td>
<td>68608</td>
<td>71715</td>
<td>-</td>
<td>1036</td>
<td>116.6</td>
<td>DNA polymerase</td>
<td>UL55 (58.4)</td>
<td>Rh89 (79.5) rhUL55 (76.8)</td>
</tr>
<tr>
<td>cy86</td>
<td>71536</td>
<td>71913</td>
<td>+</td>
<td>126</td>
<td>14.1</td>
<td>DNA polymerase</td>
<td>UL56 (73.2)</td>
<td>Rh91 (98.6) rhUL56 (98.7)</td>
</tr>
<tr>
<td>cyUL55</td>
<td>71734</td>
<td>74283</td>
<td>-</td>
<td>850</td>
<td>97.1</td>
<td>Glycoprotein B</td>
<td>UL57 (74.8)</td>
<td>rh91.1 (97) rhUL57 (97.6)</td>
</tr>
<tr>
<td>cyUL56</td>
<td>74249</td>
<td>76558</td>
<td>-</td>
<td>770</td>
<td>88.3</td>
<td>DNA packaging terminase component</td>
<td>UL58 (99.7)</td>
<td>rh93 (86.6) rhUL58 (87.8)</td>
</tr>
<tr>
<td>cy89</td>
<td>75252</td>
<td>76622</td>
<td>-</td>
<td>457</td>
<td>48.4</td>
<td>DNA packaging terminase component</td>
<td>UL59 (99.6)</td>
<td>rh92 (86.4) rhUL59 (87.6)</td>
</tr>
<tr>
<td>cyUL57</td>
<td>76705</td>
<td>80196</td>
<td>-</td>
<td>1164</td>
<td>129.3</td>
<td>Single-stranded DNA-binding protein</td>
<td>UL60 (99.6)</td>
<td>rh93 (86.6) rhUL60 (87.8)</td>
</tr>
<tr>
<td>cy91</td>
<td>81413</td>
<td>81661</td>
<td>-</td>
<td>83</td>
<td>9.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cy92</td>
<td>81579</td>
<td>81770</td>
<td>-</td>
<td>64</td>
<td>6.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>cy93</td>
<td>82075</td>
<td>82446</td>
<td>-</td>
<td>124</td>
<td>12.6</td>
<td></td>
<td>68.1</td>
<td>180.92</td>
</tr>
<tr>
<td>cyUL69</td>
<td>84231</td>
<td>86564</td>
<td>-</td>
<td>778</td>
<td>87.4</td>
<td>Tegument protein; multiple regulatory protein</td>
<td>UL69&lt;sup&gt;e&lt;/sup&gt; (49.5)</td>
<td>Rh97 (96) rhUL69 (96.1)</td>
</tr>
<tr>
<td>cy95</td>
<td>85222</td>
<td>85530</td>
<td>+</td>
<td>103</td>
<td>11.8</td>
<td></td>
<td>69.1</td>
<td>189.92</td>
</tr>
<tr>
<td>cy96</td>
<td>85773</td>
<td>86561</td>
<td>+</td>
<td>263</td>
<td>28.4</td>
<td></td>
<td>68.1</td>
<td>180.92</td>
</tr>
<tr>
<td>cyUL70</td>
<td>86498</td>
<td>89236</td>
<td>-</td>
<td>913</td>
<td>105.5</td>
<td>DNA helicase primase subunit</td>
<td>UL70 (65.3)</td>
<td>Rh100 (98.5) rhUL70 (98.5)</td>
</tr>
<tr>
<td>cy98</td>
<td>87773</td>
<td>88096</td>
<td>+</td>
<td>108</td>
<td>11.5</td>
<td></td>
<td>67.1</td>
<td>187.92</td>
</tr>
<tr>
<td>cyUL71</td>
<td>89249</td>
<td>89968</td>
<td>+</td>
<td>240</td>
<td>26.4</td>
<td>Tegument protein</td>
<td>UL71 (59.5)</td>
<td>RhUL71 (95) rhUL71 (95.4)</td>
</tr>
<tr>
<td>cyUL72</td>
<td>90036</td>
<td>91067</td>
<td>-</td>
<td>344</td>
<td>39.2</td>
<td>dUTPase/dUTPase family</td>
<td>UL72 (58.7)</td>
<td>Rh101 (98) rhUL72 (97.7)</td>
</tr>
<tr>
<td>cyUL73</td>
<td>91062</td>
<td>91373</td>
<td>+</td>
<td>104</td>
<td>11.8</td>
<td>Glycoprotein N</td>
<td>UL73 (60.8)</td>
<td>Rh102 (97.1)</td>
</tr>
<tr>
<td>cyUL74</td>
<td>91354</td>
<td>92535</td>
<td>-</td>
<td>394</td>
<td>45.9</td>
<td>Glycoprotein O</td>
<td>UL74 (43.7)</td>
<td>Rh103 (94.7) rhUL74 (94.4)</td>
</tr>
<tr>
<td>cyUL74A</td>
<td>92534</td>
<td>92704</td>
<td>+</td>
<td>57</td>
<td>6.4</td>
<td>Envelope glycoprotein 24</td>
<td>UL74A (55.3)</td>
<td></td>
</tr>
<tr>
<td>cyUL75</td>
<td>92761</td>
<td>94923</td>
<td>-</td>
<td>721</td>
<td>81.6</td>
<td>Glycoprotein H</td>
<td>UL75 (49.1)</td>
<td>Rh104 (98.2) rhUL75 (98.3)</td>
</tr>
<tr>
<td>cyUL76</td>
<td>95056</td>
<td>95940</td>
<td>+</td>
<td>295</td>
<td>32.8</td>
<td>Virion-associated regulatory protein</td>
<td>UL76 (55.4)</td>
<td>Rh105 (98.6) rhUL76 (98.3)</td>
</tr>
<tr>
<td>cyUL77</td>
<td>95609</td>
<td>97396</td>
<td>+</td>
<td>596</td>
<td>67.4</td>
<td>Portal capping protein; DNA packaging</td>
<td>UL77 (64.5)</td>
<td>Rh106 (99.5) rhUL77 (99.3)</td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>----------------</td>
<td>------------------</td>
<td>-----------------</td>
<td>--------</td>
<td>-----------</td>
<td>---------------</td>
<td>--------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>cyUL78</td>
<td>97523</td>
<td>98662</td>
<td>+</td>
<td>380</td>
<td>41.9</td>
<td>Putative chemokine receptor/ GPCR family</td>
<td>UL78 (30.4)</td>
<td>Rh107 (91.6)</td>
</tr>
<tr>
<td>cyUL79</td>
<td>98758</td>
<td>99558</td>
<td>-</td>
<td>267</td>
<td>30.5</td>
<td>Capsid maturation protease</td>
<td>UL79 (71)</td>
<td>Rh108 (98.9)</td>
</tr>
<tr>
<td>cyUL80</td>
<td>99557</td>
<td>101392</td>
<td>+</td>
<td>612</td>
<td>66.4</td>
<td>Tegument phosphoprotein pp71 (upper matrix protein)/ UL82 family; dUTPase family</td>
<td>UL80 (44.1)</td>
<td>Rh109 (97.9)</td>
</tr>
<tr>
<td>cyUL82</td>
<td>101507</td>
<td>103156</td>
<td>-</td>
<td>550</td>
<td>61.6</td>
<td>Tegument phosphoprotein pp65 (lower matrix protein)/ UL82 family; dUTPase family</td>
<td>UL82 (42.3)</td>
<td>Rh110 (95.3)</td>
</tr>
<tr>
<td>cyUL83a</td>
<td>103286</td>
<td>104911</td>
<td>-</td>
<td>542</td>
<td>62.2</td>
<td>Major tegument phosphoprotein pp65 (lower matrix protein)/ UL82 family; dUTPase family</td>
<td>UL83 (35.3)</td>
<td>Rh111 (95.6)</td>
</tr>
<tr>
<td>cy113</td>
<td>103486</td>
<td>103737</td>
<td>-</td>
<td>84</td>
<td>9.2</td>
<td></td>
<td></td>
<td>R83a (89.5)</td>
</tr>
<tr>
<td>cyUL83b</td>
<td>104980</td>
<td>106050</td>
<td>-</td>
<td>357</td>
<td>40.4</td>
<td>Major tegument phosphoprotein pp65 (lower matrix protein)/ UL82 family; dUTPase family</td>
<td>UL83 (38.9)</td>
<td>Rh112 (94.7)</td>
</tr>
<tr>
<td>cy115</td>
<td>105787</td>
<td>106218</td>
<td>+</td>
<td>144</td>
<td>15.8</td>
<td></td>
<td></td>
<td>Rh113 (87.6)</td>
</tr>
<tr>
<td>cyUL83c</td>
<td>106169</td>
<td>106594</td>
<td>-</td>
<td>142</td>
<td>15.9</td>
<td>Major tegument phosphoprotein pp65 (lower matrix protein)/ UL82 family; dUTPase family</td>
<td>UL83 (45.2)</td>
<td>Rh112 (97.1)</td>
</tr>
<tr>
<td>cyUL84</td>
<td>106714</td>
<td>108258</td>
<td>-</td>
<td>515</td>
<td>57.6</td>
<td>Role in organizing DNA replication/ UL82 family; dUTPase family</td>
<td>UL84 (50)</td>
<td>Rh114 (98.8)</td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>----------------</td>
<td>------------------</td>
<td>-----------------</td>
<td>--------</td>
<td>-----------</td>
<td>---------------</td>
<td>-------------------------------</td>
<td>--------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>cy118</td>
<td>106775</td>
<td>107149</td>
<td>+</td>
<td>125</td>
<td>13.5</td>
<td>dUTPase family</td>
<td>rh115 (96)</td>
<td>rh115 (95.2)</td>
</tr>
<tr>
<td>cy119</td>
<td>106850</td>
<td>107209</td>
<td>-</td>
<td>120</td>
<td>13.3</td>
<td></td>
<td></td>
<td>rh115.1 (90.8)</td>
</tr>
<tr>
<td>cy120</td>
<td>108000</td>
<td>108458</td>
<td>+</td>
<td>153</td>
<td>16.1</td>
<td></td>
<td>rh116 (97.4)</td>
<td></td>
</tr>
<tr>
<td>cyUL85</td>
<td>108173</td>
<td>109099</td>
<td>-</td>
<td>309</td>
<td>34.7</td>
<td>Capsid triplex subunit 2</td>
<td>UL85&lt;sup&gt;f&lt;/sup&gt; (74)</td>
<td>Rh117 (99.4)</td>
</tr>
<tr>
<td>cyUL86</td>
<td>109160</td>
<td>113191</td>
<td>-</td>
<td>1344</td>
<td>151.3</td>
<td>Major capsid protein</td>
<td>UL86 (76)</td>
<td>Rh118 (98.9)</td>
</tr>
<tr>
<td>cy123</td>
<td>109725</td>
<td>110039</td>
<td>+</td>
<td>105</td>
<td>11.8</td>
<td></td>
<td></td>
<td>rh119 (98.1)</td>
</tr>
<tr>
<td>cy124</td>
<td>110652</td>
<td>111386</td>
<td>-</td>
<td>245</td>
<td>28.7</td>
<td></td>
<td></td>
<td>rh120 (97.1)</td>
</tr>
<tr>
<td>cyUL87</td>
<td>113206</td>
<td>115758</td>
<td>+</td>
<td>851</td>
<td>96.5</td>
<td></td>
<td>UL87 (67.7)</td>
<td>Rh122 (99.4)</td>
</tr>
<tr>
<td>cyUL88</td>
<td>115771</td>
<td>116973</td>
<td>+</td>
<td>401</td>
<td>45.5</td>
<td>Tegument protein</td>
<td>UL88&lt;sup&gt;g&lt;/sup&gt; (53.7)</td>
<td>Rh123 (98)</td>
</tr>
<tr>
<td>cyUL89 ex1</td>
<td>116970</td>
<td>117917</td>
<td>-</td>
<td>316</td>
<td>35.8</td>
<td>DNA packaging terminase</td>
<td>UL89 (84.8)</td>
<td></td>
</tr>
<tr>
<td>cy128</td>
<td>117633</td>
<td>118202</td>
<td>+</td>
<td>190</td>
<td>20.8</td>
<td>component</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cyUL91</td>
<td>118234</td>
<td>118545</td>
<td>+</td>
<td>104</td>
<td>10.9</td>
<td></td>
<td>UL91 (56.3)</td>
<td>Rh126 (92.3)</td>
</tr>
<tr>
<td>cyUL92</td>
<td>118430</td>
<td>119143</td>
<td>+</td>
<td>238</td>
<td>26.5</td>
<td></td>
<td>UL92&lt;sup&gt;h&lt;/sup&gt; (90.5)</td>
<td>Rh128 (98.7)</td>
</tr>
<tr>
<td>cyUL93</td>
<td>119109</td>
<td>120671</td>
<td>+</td>
<td>521</td>
<td>59.7</td>
<td>DNA packaging; tegument</td>
<td>UL93 (48.2)</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> ORF: Open Reading Frame
<sup>b</sup> Function/gene family: dUTPase family
<sup>c</sup> HCMV: Human Cytomegalovirus
<sup>d</sup> RhCMV: Rhadinovirus Cytomegalovirus

<sup>f</sup> UL85: Capsid triplex subunit 2
<sup>g</sup> UL88: Tegument protein
<sup>h</sup> UL92: DNA packaging; tegument protein
<table>
<thead>
<tr>
<th>ORF&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Translation Start</th>
<th>Translation Stop</th>
<th>Strand</th>
<th>Size (aa)</th>
<th>Mol mass (kDa)</th>
<th>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</th>
<th>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</th>
<th>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>cyUL94</td>
<td>120547</td>
<td>121587</td>
<td>+</td>
<td>347</td>
<td>37.8</td>
<td>Tegument protein; binds single-stranded DNA</td>
<td>UL94 (58.8)</td>
<td>Rh129 (98.5)</td>
</tr>
<tr>
<td>cyUL89 ex2</td>
<td>121576</td>
<td>122508</td>
<td>-</td>
<td>311</td>
<td>36.2</td>
<td>DNA packaging terminase component</td>
<td>UL89 (86.5)</td>
<td>Rh89 (99.7)</td>
</tr>
<tr>
<td>cyUL95</td>
<td>122507</td>
<td>123802</td>
<td>+</td>
<td>432</td>
<td>47.1</td>
<td>Tegument protein</td>
<td>UL95 (66.3)</td>
<td>Rh130 (99.5)</td>
</tr>
<tr>
<td>cyUL96</td>
<td>123799</td>
<td>124188</td>
<td>+</td>
<td>130</td>
<td>14.8</td>
<td>Tegument protein</td>
<td>UL96 (59.8)</td>
<td>Rh131 (97.7)</td>
</tr>
<tr>
<td>cyUL97</td>
<td>124245</td>
<td>126071</td>
<td>+</td>
<td>609</td>
<td>67.9</td>
<td>Tegument protein; viral serine-threonine protein kinase</td>
<td>UL97 (66.5)</td>
<td>Rh132 (96.4)</td>
</tr>
<tr>
<td>cyUL98</td>
<td>126122</td>
<td>127792</td>
<td>+</td>
<td>557</td>
<td>63.4</td>
<td>Deoxyribonuclease</td>
<td>UL98 (66)</td>
<td>Rh134 (99.5)</td>
</tr>
<tr>
<td>cy138</td>
<td>126634</td>
<td>127035</td>
<td>-</td>
<td>134</td>
<td>14.6</td>
<td></td>
<td></td>
<td>Rh135 (98.5)</td>
</tr>
<tr>
<td>cy139</td>
<td>127073</td>
<td>127522</td>
<td>-</td>
<td>150</td>
<td>16.9</td>
<td></td>
<td></td>
<td>Rh136 (99.3)</td>
</tr>
<tr>
<td>cyUL99</td>
<td>127729</td>
<td>128181</td>
<td>+</td>
<td>151</td>
<td>16.5</td>
<td>Myristylated tegument phosphoprotein (pp28)</td>
<td>UL99 (57.1)</td>
<td>Rh137 (94)</td>
</tr>
<tr>
<td>cyUL100</td>
<td>128354</td>
<td>129424</td>
<td>-</td>
<td>357</td>
<td>41.0</td>
<td>Glycoprotein M</td>
<td>UL100 (58.1)</td>
<td>Rh138 (98.3)</td>
</tr>
<tr>
<td>cyUL102</td>
<td>129613</td>
<td>131790</td>
<td>+</td>
<td>726</td>
<td>80.5</td>
<td>DNA helicase primase subunit</td>
<td>UL102 (67)</td>
<td>Rh139 (97.8)</td>
</tr>
<tr>
<td>cyUL103</td>
<td>131812</td>
<td>132567</td>
<td>-</td>
<td>252</td>
<td>28.9</td>
<td>Tegument protein</td>
<td>UL103&lt;sup&gt;f&lt;/sup&gt; (55.1)</td>
<td>Rh140 (97.2)</td>
</tr>
<tr>
<td>cyUL104</td>
<td>132494</td>
<td>134464</td>
<td>-</td>
<td>657</td>
<td>75.5</td>
<td>Capsid portal protein</td>
<td>UL104 (68.3)</td>
<td>Rh141 (99.4)</td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>----------------</td>
<td>------------------</td>
<td>-----------------</td>
<td>--------</td>
<td>-----------</td>
<td>---------------</td>
<td>-------------------------------</td>
<td>---------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>cyUL105</td>
<td>134301</td>
<td>136880</td>
<td>+</td>
<td>860</td>
<td>97.6</td>
<td>DNA helicase primase subunit</td>
<td>UL105 (71.9)</td>
<td>Rh142 (99.2)</td>
</tr>
<tr>
<td>cy146</td>
<td>134450</td>
<td>134929</td>
<td>+</td>
<td>160</td>
<td>18.0</td>
<td></td>
<td></td>
<td>rh142.1 (93)</td>
</tr>
<tr>
<td>cy147</td>
<td>137750</td>
<td>137881</td>
<td>+</td>
<td>44</td>
<td>5.2</td>
<td></td>
<td></td>
<td>rh142.3 (97.6)</td>
</tr>
<tr>
<td>cy148 ex1</td>
<td>140183</td>
<td>140380</td>
<td>+</td>
<td>66</td>
<td>7.1</td>
<td>Interleukin-10-like protein precursor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cyUL111.5A ex2</td>
<td>140268</td>
<td>140729</td>
<td>+</td>
<td>154</td>
<td>17.7</td>
<td>Latency-associated viral interleukin-10</td>
<td>UL111.5A (37)</td>
<td>Rh143 (91)</td>
</tr>
<tr>
<td>cy148 ex3</td>
<td>141097</td>
<td>141186</td>
<td>+</td>
<td>30</td>
<td>3.3</td>
<td>Interleukin-10-like protein precursor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cyUL112 ex1</td>
<td>141591</td>
<td>142385</td>
<td>+</td>
<td>265</td>
<td>28.3</td>
<td>Early phosphoprotein (p50)</td>
<td>UL112 (54.3)</td>
<td>rhUL112 (91.3)</td>
</tr>
<tr>
<td>cyUL112/UL113 ex2</td>
<td>142484</td>
<td>143374</td>
<td>+</td>
<td>297</td>
<td>30.7</td>
<td>Early phosphoprotein (p84)</td>
<td>UL112/UL113 (34.8)</td>
<td>rhUL112 (94.6)</td>
</tr>
<tr>
<td>cyUL114</td>
<td>143485</td>
<td>144228</td>
<td>-</td>
<td>248</td>
<td>28.3</td>
<td>Uracil-DNA glycosylase</td>
<td>UL114 (69)</td>
<td>Rh146 (99.6)</td>
</tr>
<tr>
<td>cyUL115</td>
<td>144191</td>
<td>144967</td>
<td>-</td>
<td>259</td>
<td>29.2</td>
<td>Glycoprotein L</td>
<td>UL115 (50.6)</td>
<td>rh147 (98.8)</td>
</tr>
<tr>
<td>cyUL116</td>
<td>144978</td>
<td>146060</td>
<td>-</td>
<td>361</td>
<td>38.2</td>
<td>Putative membrane glycoprotein</td>
<td>UL116 (26.7)</td>
<td>rh148 (92.5)</td>
</tr>
<tr>
<td>cy153</td>
<td>145585</td>
<td>146082</td>
<td>+</td>
<td>166</td>
<td>16.8</td>
<td></td>
<td></td>
<td>rh147.1 (86.7)</td>
</tr>
<tr>
<td>cy154</td>
<td>145714</td>
<td>146199</td>
<td>-</td>
<td>162</td>
<td>20.1</td>
<td></td>
<td></td>
<td>rh147 (98.4)</td>
</tr>
<tr>
<td>cyUL117</td>
<td>146042</td>
<td>147190</td>
<td>-</td>
<td>383</td>
<td>42.6</td>
<td></td>
<td></td>
<td>rh149.1 (82.5)</td>
</tr>
<tr>
<td>cy156</td>
<td>147026</td>
<td>147328</td>
<td>+</td>
<td>101</td>
<td>11.1</td>
<td></td>
<td></td>
<td>rh149.1 (90.7)</td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------</td>
<td>-----------------</td>
<td>--------</td>
<td>-----------</td>
<td>---------------</td>
<td>-----------------</td>
<td>----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>cyUL119 ex1</td>
<td>147215</td>
<td>147946</td>
<td>-</td>
<td>244</td>
<td>28.4</td>
<td>Virion envelope glycoprotein; IgG Fc-binding glycoprotein</td>
<td>UL119 (35.1)</td>
<td>rhUL119 (96.5)</td>
</tr>
<tr>
<td>cy157 ex2</td>
<td>147894</td>
<td>148532</td>
<td>-</td>
<td>213</td>
<td>21.3</td>
<td></td>
<td></td>
<td>rhUL119 (79.8)</td>
</tr>
<tr>
<td>cy158</td>
<td>148043</td>
<td>148195</td>
<td>-</td>
<td>51</td>
<td>5.4</td>
<td></td>
<td></td>
<td>rh153 (76.9)</td>
</tr>
<tr>
<td>cyUL120</td>
<td>148581</td>
<td>149177</td>
<td>-</td>
<td>199</td>
<td>22.6</td>
<td>Putative membrane glycoprotein/ UL120 family</td>
<td>UL120 (44.1)</td>
<td>Rh154 (92.4)</td>
</tr>
<tr>
<td>cyUL121</td>
<td>149179</td>
<td>149727</td>
<td>-</td>
<td>183</td>
<td>21.0</td>
<td>Putative membrane glycoprotein/ UL120 family</td>
<td>UL121 (26.9)</td>
<td>Rh155 (96.7)</td>
</tr>
<tr>
<td>cyUL122 ex1</td>
<td>149990</td>
<td>151474</td>
<td>-</td>
<td>495</td>
<td>53.6</td>
<td>Immediate early 2 transactivator</td>
<td>U1L122 (58.7)</td>
<td>rhUL122 (89.9)</td>
</tr>
<tr>
<td>cyUL123 ex2</td>
<td>151977</td>
<td>153122</td>
<td>-</td>
<td>382</td>
<td>43.1</td>
<td>Majoy immediate early 1 co-transactivator</td>
<td>U1L123 (23.7)</td>
<td>rhUL123 (60.7)</td>
</tr>
<tr>
<td>cy161 ex3</td>
<td>153182</td>
<td>153559</td>
<td>-</td>
<td>126</td>
<td>14.4</td>
<td>Immediate early protein</td>
<td>U1L126 (55.8)</td>
<td>rhUL123 (84.6)</td>
</tr>
<tr>
<td>cy161 ex4</td>
<td>153665</td>
<td>153784</td>
<td>-</td>
<td>40</td>
<td>4.1</td>
<td>Immediate early protein</td>
<td></td>
<td>rh156.1 (90.8)</td>
</tr>
<tr>
<td>cy162</td>
<td>153674</td>
<td>153937</td>
<td>+</td>
<td>88</td>
<td>10.4</td>
<td></td>
<td>rh156.1 (92)</td>
<td>rh156.2 (94)</td>
</tr>
<tr>
<td>cy163</td>
<td>153783</td>
<td>154238</td>
<td>+</td>
<td>152</td>
<td>15.8</td>
<td></td>
<td>rh156.2 (93)</td>
<td>rh157.1 (61)</td>
</tr>
<tr>
<td>cyUL126</td>
<td>154779</td>
<td>154925</td>
<td>-</td>
<td>49</td>
<td>5.7</td>
<td></td>
<td>UL126 (55.8)</td>
<td>rh157.1 (87)</td>
</tr>
<tr>
<td>cy165</td>
<td>155301</td>
<td>155795</td>
<td>-</td>
<td>165</td>
<td>18.6</td>
<td></td>
<td>rh157.1 (87)</td>
<td>rh157.1 (87.8)</td>
</tr>
<tr>
<td>cy166</td>
<td>155342</td>
<td>155491</td>
<td>+</td>
<td>50</td>
<td>5.8</td>
<td></td>
<td>rh157.1 (61)</td>
<td>rh157.1 (61.4)</td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>----------------</td>
<td>------------------</td>
<td>------------------</td>
<td>--------</td>
<td>-----------</td>
<td>---------------</td>
<td>---------------------------------</td>
<td>----------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>cy167</td>
<td>155476</td>
<td>155679</td>
<td>+</td>
<td>68</td>
<td>7.1</td>
<td></td>
<td>68.1</td>
<td>180.92</td>
</tr>
<tr>
<td>cy168</td>
<td>155547</td>
<td>155780</td>
<td>+</td>
<td>78</td>
<td>8.4</td>
<td></td>
<td>rh157.3</td>
<td>55.1</td>
</tr>
<tr>
<td>cy169</td>
<td>155555</td>
<td>155704</td>
<td>-</td>
<td>50</td>
<td>5.2</td>
<td></td>
<td>rh157.1</td>
<td>53.6</td>
</tr>
<tr>
<td>cy170</td>
<td>155717</td>
<td>156094</td>
<td>-</td>
<td>126</td>
<td>14.8</td>
<td></td>
<td>rh157</td>
<td>91.8</td>
</tr>
<tr>
<td>cy171</td>
<td>155719</td>
<td>156183</td>
<td>+</td>
<td>155</td>
<td>18.3</td>
<td></td>
<td>rh157.1</td>
<td>79.6</td>
</tr>
<tr>
<td>cy172</td>
<td>155823</td>
<td>156128</td>
<td>+</td>
<td>102</td>
<td>11.6</td>
<td></td>
<td>rh157.2</td>
<td>80.9</td>
</tr>
<tr>
<td>cyUL128 ex1</td>
<td>156700</td>
<td>157182</td>
<td>-</td>
<td>161</td>
<td>18.3</td>
<td>Putative secreted protein; putative CC chemokine</td>
<td>UL128 (37.1)</td>
<td>rhUL128 (96.7)</td>
</tr>
<tr>
<td>cyUL128 ex2</td>
<td>157420</td>
<td>157608</td>
<td>-</td>
<td>63</td>
<td>7.1</td>
<td>Putative secreted protein; putative CC chemokine</td>
<td>UL128 (55.2)</td>
<td>rhUL128 (93.6)</td>
</tr>
<tr>
<td>cy174</td>
<td>157610</td>
<td>158317</td>
<td>-</td>
<td>236</td>
<td>25.5</td>
<td></td>
<td>rh157.4</td>
<td>76.2</td>
</tr>
<tr>
<td>cyUL130</td>
<td>158438</td>
<td>158731</td>
<td>-</td>
<td>98</td>
<td>11.5</td>
<td>Putative secreted protein</td>
<td>UL130 (40.7)</td>
<td>rhUL130 (85.6)</td>
</tr>
<tr>
<td>cyUL131A</td>
<td>159065</td>
<td>159322</td>
<td>-</td>
<td>86</td>
<td>9.7</td>
<td>Putative secreted protein</td>
<td>UL131A (35.3)</td>
<td>rh131a (96.5)</td>
</tr>
<tr>
<td>cyUL132</td>
<td>159366</td>
<td>160025</td>
<td>-</td>
<td>220</td>
<td>24.1</td>
<td>Envelope glycoprotein</td>
<td>UL132 (32.1)</td>
<td>Rh160 (94.9)</td>
</tr>
<tr>
<td>cyUL148</td>
<td>160091</td>
<td>161071</td>
<td>-</td>
<td>327</td>
<td>37.0</td>
<td>Putative membrane glycoprotein</td>
<td>UL148 (30.5)</td>
<td>Rh159 (89)</td>
</tr>
<tr>
<td>cyUL147</td>
<td>161273</td>
<td>161707</td>
<td>-</td>
<td>145</td>
<td>16.5</td>
<td>Chemokine vCXCL2/ UL146 family</td>
<td>UL147 (43.3)</td>
<td>Rh158 (88.2)</td>
</tr>
<tr>
<td>cyUL146</td>
<td>161786</td>
<td>162148</td>
<td>-</td>
<td>121</td>
<td>13.5</td>
<td>Chemokine vCXCL1/ UL146</td>
<td>UL146</td>
<td></td>
</tr>
<tr>
<td>ORFa</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene familyb</td>
<td>HCMVc Homologue (%)d</td>
<td>RhCMV Homologue (%)d</td>
</tr>
<tr>
<td>------</td>
<td>------------------</td>
<td>------------------</td>
<td>-------</td>
<td>----------</td>
<td>---------------</td>
<td>-----------------------</td>
<td>-----------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>cy181</td>
<td>162372</td>
<td>162590</td>
<td>-</td>
<td>73</td>
<td>8.2</td>
<td>Alpha-chemokine-like protein</td>
<td></td>
<td>68.1</td>
</tr>
<tr>
<td>cy182</td>
<td>162689</td>
<td>163015</td>
<td>-</td>
<td>109</td>
<td>12.1</td>
<td>Alpha-chemokine-like protein</td>
<td></td>
<td>180.92</td>
</tr>
<tr>
<td>cy183</td>
<td>163138</td>
<td>163470</td>
<td>-</td>
<td>111</td>
<td>12.5</td>
<td>Alpha-chemokine-like protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cy184</td>
<td>163554</td>
<td>164072</td>
<td>-</td>
<td>173</td>
<td>19.3</td>
<td>Alpha-chemokine-like protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cyUL145</td>
<td>164157</td>
<td>164462</td>
<td>-</td>
<td>102</td>
<td>11.3</td>
<td>Membrane glycoprotein; tumor necrosis factor-receptor homologue</td>
<td>UL145 (65.6)</td>
<td>Rh162 (99)</td>
</tr>
<tr>
<td>cyUL144</td>
<td>164880</td>
<td>165395</td>
<td>-</td>
<td>172</td>
<td>18.7</td>
<td>Membrane glycoprotein; tumor necrosis factor-receptor homologue</td>
<td>UL144f (29.6)</td>
<td>Rh163 (98.8)</td>
</tr>
<tr>
<td>cyUL141</td>
<td>165611</td>
<td>166909</td>
<td>-</td>
<td>433</td>
<td>49.1</td>
<td>Membrane glycoprotein/UL14 family</td>
<td>UL141 (39.4)</td>
<td>Rh164 (97)</td>
</tr>
<tr>
<td>cy188</td>
<td>167408</td>
<td>167854</td>
<td>-</td>
<td>149</td>
<td>17.1</td>
<td>Membrane glycoprotein/UL14 family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cy189</td>
<td>167899</td>
<td>168426</td>
<td>-</td>
<td>176</td>
<td>19.4</td>
<td>Membrane glycoprotein/UL14 family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cy190</td>
<td>168557</td>
<td>169060</td>
<td>-</td>
<td>168</td>
<td>18.2</td>
<td>Membrane glycoprotein/UL14 family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cy191</td>
<td>169318</td>
<td>169977</td>
<td>-</td>
<td>220</td>
<td>24.7</td>
<td>Membrane glycoprotein/UL14 family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cy192</td>
<td>170074</td>
<td>170505</td>
<td>+</td>
<td>144</td>
<td>16.3</td>
<td>Membrane glycoprotein/UL14 family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cy193</td>
<td>170080</td>
<td>170643</td>
<td>-</td>
<td>188</td>
<td>20.8</td>
<td>Membrane glycoprotein/UL14 family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>cy194</td>
<td>170777</td>
<td>171343</td>
<td>-</td>
<td>189</td>
<td>21.2</td>
<td></td>
<td>rh170 (96.8)</td>
<td>rh170 (97.3)</td>
</tr>
<tr>
<td>cy195</td>
<td>171346</td>
<td>172185</td>
<td>-</td>
<td>280</td>
<td>30.7</td>
<td></td>
<td>Rh171 (91.4)</td>
<td>rh171 (90.7)</td>
</tr>
<tr>
<td>cy196</td>
<td>172385</td>
<td>172915</td>
<td>-</td>
<td>177</td>
<td>19.9</td>
<td></td>
<td>rh172 (96)</td>
<td>rh172 (97.2)</td>
</tr>
<tr>
<td>cy197</td>
<td>172418</td>
<td>172792</td>
<td>+</td>
<td>125</td>
<td>13.9</td>
<td></td>
<td>rh171.1 (90)</td>
<td>rh171.1 (91.9)</td>
</tr>
<tr>
<td>cyUL153</td>
<td>172967</td>
<td>174088</td>
<td>-</td>
<td>374</td>
<td>40.6</td>
<td>Membrane protein RL13/RL11 family</td>
<td>UL153&lt;sup&gt;h&lt;/sup&gt; (32.8)</td>
<td>rh173 (57.3)</td>
</tr>
<tr>
<td>cy199</td>
<td>175240</td>
<td>176325</td>
<td>-</td>
<td>362</td>
<td>39.9</td>
<td></td>
<td>rh174 (92.2)</td>
<td>rh174 (92.2)</td>
</tr>
<tr>
<td>cy200</td>
<td>177022</td>
<td>177480</td>
<td>+</td>
<td>153</td>
<td>16.5</td>
<td></td>
<td>rh175 (92.8)</td>
<td>rh175 (93.4)</td>
</tr>
<tr>
<td>cy201</td>
<td>177171</td>
<td>177824</td>
<td>-</td>
<td>218</td>
<td>23.7</td>
<td></td>
<td>rh176 (92.2)</td>
<td>rh176 (93.1)</td>
</tr>
<tr>
<td>cy202</td>
<td>177790</td>
<td>178263</td>
<td>-</td>
<td>158</td>
<td>17.8</td>
<td></td>
<td>rh177 (78.8)</td>
<td>rh177 (78.2)</td>
</tr>
<tr>
<td>cy203</td>
<td>177884</td>
<td>178657</td>
<td>-</td>
<td>258</td>
<td>28.2</td>
<td></td>
<td>rh178 (86.1)</td>
<td>rh178 (84.9)</td>
</tr>
<tr>
<td>cy204</td>
<td>178734</td>
<td>178931</td>
<td>-</td>
<td>66</td>
<td>7.0</td>
<td></td>
<td>rh178.2 (82.8)</td>
<td>rh178.2 (82.8)</td>
</tr>
<tr>
<td>cy205</td>
<td>178882</td>
<td>179079</td>
<td>+</td>
<td>66</td>
<td>6.9</td>
<td></td>
<td>rh178.1 (51)</td>
<td>rh178.1 (47.8)</td>
</tr>
<tr>
<td>cy206</td>
<td>179072</td>
<td>179575</td>
<td>+</td>
<td>168</td>
<td>18.2</td>
<td></td>
<td>rh178.3 (64.1)</td>
<td>rh178.3 (71.7)</td>
</tr>
<tr>
<td>cy207</td>
<td>179212</td>
<td>179535</td>
<td>+</td>
<td>108</td>
<td>11.8</td>
<td></td>
<td>rh178.3 (78)</td>
<td>rh178.3 (80.8)</td>
</tr>
<tr>
<td>cy208</td>
<td>179835</td>
<td>180350</td>
<td>+</td>
<td>172</td>
<td>18.6</td>
<td></td>
<td>rh179</td>
<td>rh179</td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------</td>
<td>------------------</td>
<td>--------</td>
<td>-----------</td>
<td>----------------</td>
<td>-------------------------------</td>
<td>-----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>cy209</td>
<td>179951</td>
<td>180157</td>
<td>-</td>
<td>69</td>
<td>6.8</td>
<td></td>
<td></td>
<td>(92.4)</td>
</tr>
<tr>
<td>cyUS1</td>
<td>180377</td>
<td>180883</td>
<td>-</td>
<td>169</td>
<td>19.3</td>
<td>US1 family</td>
<td>US1</td>
<td>(49.4)</td>
</tr>
<tr>
<td>cy211</td>
<td>180732</td>
<td>180965</td>
<td>+</td>
<td>78</td>
<td>8.3</td>
<td></td>
<td></td>
<td>(91)</td>
</tr>
<tr>
<td>cyUS2</td>
<td>181118</td>
<td>181708</td>
<td>-</td>
<td>197</td>
<td>23.3</td>
<td>Membrane glycoprotein/ US2 family</td>
<td>US2</td>
<td>(21)</td>
</tr>
<tr>
<td>cyUS3</td>
<td>182241</td>
<td>182786</td>
<td>-</td>
<td>182</td>
<td>20.7</td>
<td>Immediate early glycoprotein/ US2 family</td>
<td>US3</td>
<td>(26.4)</td>
</tr>
<tr>
<td>cy214</td>
<td>182286</td>
<td>182579</td>
<td>+</td>
<td>98</td>
<td>10.8</td>
<td></td>
<td></td>
<td>(90.7)</td>
</tr>
<tr>
<td>cy215</td>
<td>183413</td>
<td>183712</td>
<td>-</td>
<td>100</td>
<td>11.1</td>
<td></td>
<td></td>
<td>rh184.1</td>
</tr>
<tr>
<td>cy216</td>
<td>183794</td>
<td>184366</td>
<td>-</td>
<td>191</td>
<td>21.4</td>
<td></td>
<td></td>
<td>rh185</td>
</tr>
<tr>
<td>cy217</td>
<td>184596</td>
<td>185300</td>
<td>-</td>
<td>235</td>
<td>27.8</td>
<td></td>
<td></td>
<td>rh186</td>
</tr>
<tr>
<td>cyUS11a</td>
<td>185545</td>
<td>186228</td>
<td>-</td>
<td>228</td>
<td>25.7</td>
<td>Membrane glycoprotein/ US6 family</td>
<td>US11</td>
<td>(24)</td>
</tr>
<tr>
<td>cy219</td>
<td>186324</td>
<td>186698</td>
<td>-</td>
<td>125</td>
<td>14.6</td>
<td></td>
<td></td>
<td>rh188</td>
</tr>
<tr>
<td>cyUS11b</td>
<td>186981</td>
<td>187823</td>
<td>-</td>
<td>281</td>
<td>32.8</td>
<td>Membrane glycoprotein/ US6 family</td>
<td>US11</td>
<td>(29.1)</td>
</tr>
<tr>
<td>cyUS12</td>
<td>188025</td>
<td>188807</td>
<td>-</td>
<td>261</td>
<td>30.0</td>
<td>Putative multiple transmembrane protein/ US12 family</td>
<td>US12</td>
<td>(37.2)</td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>cy222</td>
<td>188182</td>
<td>188334</td>
<td>-</td>
<td>51</td>
<td>5.4</td>
<td></td>
<td>68.1</td>
<td>180.92</td>
</tr>
<tr>
<td>cy223</td>
<td>188395</td>
<td>188724</td>
<td>-</td>
<td>110</td>
<td>12.0</td>
<td></td>
<td>rh191 (84)</td>
<td>rh191 (87.2)</td>
</tr>
<tr>
<td>cyUS13</td>
<td>188865</td>
<td>189629</td>
<td>-</td>
<td>255</td>
<td>29.7</td>
<td>Putative multiple transmembrane protein/ US12 family</td>
<td>US13&lt;sup&gt;f&lt;/sup&gt; (25.9)</td>
<td>Rh192 (99.6) rhUS13 (99.2)</td>
</tr>
<tr>
<td>cyUS14a</td>
<td>189742</td>
<td>190575</td>
<td>-</td>
<td>278</td>
<td>31.4</td>
<td>Putative multiple transmembrane protein/ US12 family</td>
<td>US14 (25.3)</td>
<td>Rh194 (97.8) rh194 (98.2)</td>
</tr>
<tr>
<td>cyUS14b</td>
<td>190706</td>
<td>191434</td>
<td>-</td>
<td>243</td>
<td>27.4</td>
<td>Putative multiple transmembrane protein/ US12 family</td>
<td>US14 (25.3)</td>
<td>Rh195 (98.8) rhUS14 (26.4)</td>
</tr>
<tr>
<td>cyUS14c</td>
<td>191527</td>
<td>192285</td>
<td>-</td>
<td>253</td>
<td>29.5</td>
<td>Putative multiple transmembrane protein/ US12 family</td>
<td>US14 (29.7)</td>
<td>Rh196 (98.4) rhUS14 (98.8)</td>
</tr>
<tr>
<td>cy228</td>
<td>192391</td>
<td>193116</td>
<td>-</td>
<td>242</td>
<td>28.0</td>
<td></td>
<td>Rh197 (96.3)</td>
<td>rh197 (96.3) rh196.1 (92.1)</td>
</tr>
<tr>
<td>cy229</td>
<td>192716</td>
<td>192985</td>
<td>-</td>
<td>90</td>
<td>10.8</td>
<td></td>
<td></td>
<td>rh198 (97.4) rh197 (98.1)</td>
</tr>
<tr>
<td>cyUS17</td>
<td>193094</td>
<td>193918</td>
<td>-</td>
<td>275</td>
<td>30.4</td>
<td>Putative multiple transmembrane protein/ US12 family</td>
<td>US17 (42.9)</td>
<td>Rh198 (98.2) rhUS17 (97.4)</td>
</tr>
<tr>
<td>cyUS18</td>
<td>194024</td>
<td>194824</td>
<td>-</td>
<td>267</td>
<td>30.1</td>
<td>Putative multiple transmembrane protein/ US12 family</td>
<td>US18&lt;sup&gt;f&lt;/sup&gt; (28.6)</td>
<td>rhUS18 (98.1)</td>
</tr>
<tr>
<td>cyUS19</td>
<td>194944</td>
<td>195729</td>
<td>-</td>
<td>262</td>
<td>30.0</td>
<td>Putative multiple transmembrane protein/ US12 family</td>
<td>US19 (27.8)</td>
<td>Rh200 (95) rh200 (95)</td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>----------------</td>
<td>------------------</td>
<td>-----------------</td>
<td>--------</td>
<td>-----------</td>
<td>----------------</td>
<td>-------------------------------</td>
<td>---------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>cyUS20</td>
<td>195790</td>
<td>196551</td>
<td>-</td>
<td>254</td>
<td>28.6</td>
<td>Putative multiple transmembrane protein/ US12 family</td>
<td>US20 (43.9)</td>
<td>Rh201 (99.6)</td>
</tr>
<tr>
<td>cyUS21</td>
<td>196599</td>
<td>197285</td>
<td>-</td>
<td>229</td>
<td>26.1</td>
<td>Putative multiple transmembrane protein/ US12 family</td>
<td>US21 (61.1)</td>
<td>Rh202 (98.2)</td>
</tr>
<tr>
<td>cyUS22</td>
<td>197407</td>
<td>199131</td>
<td>-</td>
<td>575</td>
<td>65.8</td>
<td>Tegument protein/ US22 family</td>
<td>US22 (46.3)</td>
<td>Rh203 (96.9)</td>
</tr>
<tr>
<td>cyUS23</td>
<td>199290</td>
<td>201161</td>
<td>-</td>
<td>624</td>
<td>72.7</td>
<td>Tegument protein/ US22 family</td>
<td>US23 (51.8)</td>
<td>Rh204 (98.2)</td>
</tr>
<tr>
<td>cy237</td>
<td>199589</td>
<td>200002</td>
<td>-</td>
<td>138</td>
<td>16.5</td>
<td></td>
<td></td>
<td>rh206 (90.5)</td>
</tr>
<tr>
<td>cy238</td>
<td>199788</td>
<td>200117</td>
<td>+</td>
<td>110</td>
<td>12.9</td>
<td></td>
<td></td>
<td>rh207 (92.5)</td>
</tr>
<tr>
<td>cy239</td>
<td>200946</td>
<td>201263</td>
<td>+</td>
<td>106</td>
<td>11.5</td>
<td></td>
<td></td>
<td>rh208 (89.5)</td>
</tr>
<tr>
<td>cyUS24</td>
<td>201185</td>
<td>202615</td>
<td>-</td>
<td>477</td>
<td>56.6</td>
<td>Tegument protein/ US22 family</td>
<td>US24 (66.2)</td>
<td>rh209 (98.7)</td>
</tr>
<tr>
<td>cy241</td>
<td>202022</td>
<td>202207</td>
<td>+</td>
<td>62</td>
<td>7.0</td>
<td></td>
<td></td>
<td>rh210 (90)</td>
</tr>
<tr>
<td>cyUS26</td>
<td>202978</td>
<td>204759</td>
<td>-</td>
<td>594</td>
<td>67.3</td>
<td>US22 family</td>
<td>US26 (46.8)</td>
<td>rh211 (96.8)</td>
</tr>
<tr>
<td>cy243</td>
<td>203421</td>
<td>203735</td>
<td>-</td>
<td>105</td>
<td>11.6</td>
<td></td>
<td></td>
<td>rh212 (98.1)</td>
</tr>
<tr>
<td>cy244</td>
<td>203512</td>
<td>204027</td>
<td>+</td>
<td>172</td>
<td>19.0</td>
<td></td>
<td></td>
<td>rh214 (94.7)</td>
</tr>
<tr>
<td>cyUS28a</td>
<td>204931</td>
<td>205917</td>
<td>+</td>
<td>329</td>
<td>37.5</td>
<td>Membrane protein; CC and CX3C chemokine receptor; mediates cellular activation</td>
<td>US28 (29.1)</td>
<td>Rh214 (97.9)</td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Translation Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------</td>
<td>-----------------</td>
<td>--------</td>
<td>-----------</td>
<td>---------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>cyUS28b</td>
<td>206261</td>
<td>207274</td>
<td>+</td>
<td>338</td>
<td>38.7</td>
<td>Membrane protein; CC and CX3C chemokine receptor; mediates cellular activation and migration; virion envelope glycoprotein/ GPCR family</td>
<td>US28 (25.4)</td>
<td>Rh215 (93.5)</td>
</tr>
<tr>
<td>cyUS28c</td>
<td>207400</td>
<td>208401</td>
<td>+</td>
<td>334</td>
<td>38.1</td>
<td>Membrane protein; CC and CX3C chemokine receptor; mediates cellular activation and migration; virion envelope glycoprotein/ GPCR family</td>
<td>US28 (24.5)</td>
<td>rhUS28.2 (97.3)</td>
</tr>
<tr>
<td>cy248</td>
<td>208296</td>
<td>208568</td>
<td>-</td>
<td>91</td>
<td>10.2</td>
<td>Membrane protein; CC and CX3C chemokine receptor; mediates cellular activation and migration; virion envelope glycoprotein/ GPCR family</td>
<td></td>
<td>Rh217 (88.9)</td>
</tr>
<tr>
<td>cyUS28d</td>
<td>208474</td>
<td>209496</td>
<td>+</td>
<td>341</td>
<td>39.1</td>
<td>Membrane protein; CC and CX3C chemokine receptor; mediates cellular activation and migration; virion envelope glycoprotein/ GPCR family</td>
<td>US28 (26.2)</td>
<td>Rh218 (96.2)</td>
</tr>
<tr>
<td>cy250</td>
<td>209266</td>
<td>209571</td>
<td>-</td>
<td>102</td>
<td>11.4</td>
<td>Membrane protein; CC and CX3C chemokine receptor; mediates cellular activation and migration; virion envelope glycoprotein/ GPCR family</td>
<td></td>
<td>rh219 (94.1)</td>
</tr>
<tr>
<td>cyUS28e</td>
<td>209641</td>
<td>211104</td>
<td>+</td>
<td>488</td>
<td>54.0</td>
<td>Membrane protein; CC and CX3C chemokine receptor; mediates cellular activation and migration; virion envelope glycoprotein/ GPCR family</td>
<td>US28 (40)</td>
<td>Rh220 (85.3)</td>
</tr>
<tr>
<td>ORF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Translation Start</td>
<td>Stop</td>
<td>Strand</td>
<td>Size (aa)</td>
<td>Mol mass (kDa)</td>
<td>Function/gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>HCMV&lt;sup&gt;c&lt;/sup&gt; Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>RhCMV Homologue (%)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>cyUS29</td>
<td>211265</td>
<td>212581</td>
<td>+</td>
<td>439</td>
<td>49.3</td>
<td>Putative membrane glycoprotein</td>
<td>US29 (38.1)</td>
<td>Rh221 (95)</td>
</tr>
<tr>
<td>cy253</td>
<td>211728</td>
<td>212051</td>
<td>+</td>
<td>108</td>
<td>12.7</td>
<td></td>
<td></td>
<td>rh222 (97.2)</td>
</tr>
<tr>
<td>cyUS30</td>
<td>212499</td>
<td>213320</td>
<td>+</td>
<td>274</td>
<td>30.8</td>
<td>Putative membrane glycoprotein</td>
<td>US30 (22.2)</td>
<td>Rh223 (96)</td>
</tr>
<tr>
<td>cy255</td>
<td>213182</td>
<td>213544</td>
<td>-</td>
<td>121</td>
<td>12.6</td>
<td></td>
<td></td>
<td>rh224 (95.8)</td>
</tr>
<tr>
<td>cyUS31</td>
<td>213396</td>
<td>213881</td>
<td>+</td>
<td>162</td>
<td>18.5</td>
<td>US1 family</td>
<td>US31 (42.7)</td>
<td>Rh225 (94.4)</td>
</tr>
<tr>
<td>cy257</td>
<td>213572</td>
<td>213799</td>
<td>-</td>
<td>76</td>
<td>8.5</td>
<td></td>
<td></td>
<td>rh224 (95.2)</td>
</tr>
<tr>
<td>cyUS32</td>
<td>214008</td>
<td>214568</td>
<td>+</td>
<td>187</td>
<td>22.2</td>
<td>US1 family</td>
<td>US32 (44.7)</td>
<td>Rh226 (96.8)</td>
</tr>
<tr>
<td>cy259</td>
<td>214288</td>
<td>214482</td>
<td>-</td>
<td>65</td>
<td>7.2</td>
<td></td>
<td></td>
<td>rh227 (93.7)</td>
</tr>
<tr>
<td>cy260</td>
<td>214712</td>
<td>215017</td>
<td>+</td>
<td>102</td>
<td>10.9</td>
<td></td>
<td></td>
<td>rh228 (93.1)</td>
</tr>
<tr>
<td>cy261</td>
<td>215664</td>
<td>216176</td>
<td>-</td>
<td>171</td>
<td>18.6</td>
<td></td>
<td></td>
<td>rh229 (86.2)</td>
</tr>
<tr>
<td>cyTRS1</td>
<td>215743</td>
<td>217818</td>
<td>-</td>
<td>692</td>
<td>77.3</td>
<td>Tegument protein; immediate early protein/ US22 family</td>
<td>TRS1 (37.4)</td>
<td>Rh230 (92.4)</td>
</tr>
</tbody>
</table>

<sup>a</sup> The CyCMV genes with an HCMV homologue are annotated by “cy” followed by the HCMV name.

<sup>b</sup> Functions and gene families were assigned based on studies of HCMV (553).

<sup>c</sup> HCMV homologues are from strains AD169 and Toledo unless otherwise specified.
\(^d\) Percent identity based on BLASTP search conducted in June 2011.

\(^e\) HCMV strain 3301.

\(^f\) HCMV strain Merlin.

\(^g\) HCMV strain CINCY and Towne.

\(^h\) HCMV strain Towne.

\(^i\) HCMV strain NT.
<table>
<thead>
<tr>
<th>Missing HCMV Genes&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Function/ gene family&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Effect of deletion on viral growth kinetics&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>RL2</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>RL3</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>RL4</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>RL5</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>RL6</td>
<td>RL11 family</td>
<td>-</td>
</tr>
<tr>
<td>RL7</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>RL8</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>RL9</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>RL10</td>
<td>Envelope glycoprotein</td>
<td>-</td>
</tr>
<tr>
<td>RL12</td>
<td>Putative membrane glycoprotein/ RL11 family</td>
<td>-</td>
</tr>
<tr>
<td>RL13</td>
<td>Putative membrane glycoprotein/ RL11 family</td>
<td>-</td>
</tr>
<tr>
<td>RL14</td>
<td>RL11 family</td>
<td>ND</td>
</tr>
<tr>
<td>UL1</td>
<td>RL11 family</td>
<td>ND</td>
</tr>
<tr>
<td>UL2</td>
<td>Putative MP</td>
<td>+</td>
</tr>
<tr>
<td>UL3</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>UL4</td>
<td>Transcriptionally regulated envelope glycoprotein/ RL11 family</td>
<td>-</td>
</tr>
<tr>
<td>UL5</td>
<td>RL11 family</td>
<td>-</td>
</tr>
<tr>
<td>UL8</td>
<td>RL11 family</td>
<td>-</td>
</tr>
<tr>
<td>UL10</td>
<td>Putative membrane glycoprotein/ RL11 family</td>
<td>-</td>
</tr>
<tr>
<td>UL12</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>UL15</td>
<td>Putative MP</td>
<td>-</td>
</tr>
<tr>
<td>Missing HCMV Genes (^a)</td>
<td>Function/ gene family (^b)</td>
<td>Effect of deletion on viral growth kinetics (^c)</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>UL16</td>
<td>Membrane glycoprotein</td>
<td>-</td>
</tr>
<tr>
<td>UL17</td>
<td>Seven-TM membrane glycoprotein</td>
<td>-</td>
</tr>
<tr>
<td>UL18</td>
<td>Putative membrane glycoprotein; MHC-I homologue/ UL18 family</td>
<td>-</td>
</tr>
<tr>
<td>UL22</td>
<td>Envelope glycoprotein; secreted glycoprotein</td>
<td>ND</td>
</tr>
<tr>
<td>UL39</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>UL40</td>
<td>Membrane glycoprotein</td>
<td>ND</td>
</tr>
<tr>
<td>UL58</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>UL59</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>UL60</td>
<td></td>
<td>++++</td>
</tr>
<tr>
<td>UL61</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>UL62</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>UL63</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>UL64</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>UL65</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>UL66</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>UL67</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>UL68</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>UL80.5</td>
<td>Capsid scaffold protein</td>
<td>ND</td>
</tr>
<tr>
<td>UL81</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>UL90</td>
<td></td>
<td>++++</td>
</tr>
<tr>
<td>UL101</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>UL106</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>UL107</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>Missing HCMV Genes&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Function/ gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Effect of deletion on viral growth kinetics&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>UL108</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>UL109</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>UL110</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>UL118</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>UL124</td>
<td>Putative membrane glycoprotein</td>
<td>+ +&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>UL125</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>UL127</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>UL129</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>UL147A</td>
<td>Putative MP</td>
<td>ND</td>
</tr>
<tr>
<td>UL143</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>UL142</td>
<td>Putative membrane glycoprotein; MHC-I homologue/ UL18 family</td>
<td>ND</td>
</tr>
<tr>
<td>UL140</td>
<td>Putative MP</td>
<td>ND</td>
</tr>
<tr>
<td>UL139</td>
<td>Putative membrane glycoprotein</td>
<td>ND</td>
</tr>
<tr>
<td>UL138</td>
<td>Putative MP</td>
<td>ND</td>
</tr>
<tr>
<td>UL137</td>
<td>Putative MP</td>
<td>ND</td>
</tr>
<tr>
<td>UL136</td>
<td>Putative MP</td>
<td>ND</td>
</tr>
<tr>
<td>UL135</td>
<td>Putative secreted protein</td>
<td>ND</td>
</tr>
<tr>
<td>UL134</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>UL133</td>
<td>Putative MP</td>
<td>ND</td>
</tr>
<tr>
<td>UL148A</td>
<td>Putative MP</td>
<td>ND</td>
</tr>
<tr>
<td>UL148B</td>
<td>Putative MP</td>
<td>ND</td>
</tr>
<tr>
<td>UL148C</td>
<td>Putative MP</td>
<td>ND</td>
</tr>
<tr>
<td>UL148D</td>
<td>Putative MP</td>
<td>ND</td>
</tr>
<tr>
<td>Missing HCMV Genes&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Function/ gene family&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Effect of deletion on viral growth kinetics&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>UL149</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>UL150</td>
<td>Putative secreted protein</td>
<td>ND</td>
</tr>
<tr>
<td>IRS1</td>
<td>Immediate-early protein; TP/US22 family</td>
<td></td>
</tr>
<tr>
<td>US4</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>US5</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>US6</td>
<td>Putative membrane glycoprotein/ US6 family</td>
<td></td>
</tr>
<tr>
<td>US7</td>
<td>Membrane glycoprotein/ US6 family</td>
<td></td>
</tr>
<tr>
<td>US8</td>
<td>Membrane glycoprotein/ US6 family</td>
<td></td>
</tr>
<tr>
<td>US9</td>
<td>Membrane glycoprotein/ US6 family</td>
<td></td>
</tr>
<tr>
<td>US10</td>
<td>Membrane glycoprotein/ US6 family</td>
<td></td>
</tr>
<tr>
<td>US15</td>
<td>Putative multiple-transmembrane protein/ US12 family</td>
<td></td>
</tr>
<tr>
<td>US16</td>
<td>Putative multiple-transmembrane protein/ US12 family</td>
<td></td>
</tr>
<tr>
<td>US25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US27</td>
<td>Virion envelope glycoprotein/GPCR family</td>
<td></td>
</tr>
<tr>
<td>US33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US34</td>
<td>Putative secreted protein</td>
<td></td>
</tr>
<tr>
<td>US34A</td>
<td>Putative MP</td>
<td>ND</td>
</tr>
</tbody>
</table>

<sup>a</sup> According to ORFs of HCMV strains AD169 (X17403) and Toledo (GU937742).

<sup>b</sup> Functions and gene families were assigned based on studies of HCMV (553).

<sup>c</sup> ND, not determined; -, no effect; +, modest effect on viral replication; ++, critical effect on viral replication; +++ , required for viral replication.

<sup>d</sup> Effects of deletion differ between studies (233).
Chapter 4

4 The Development of Cynomolgus Macaque Cytomegalovirus (CyCMV) as a Viral Vector for an HIV/SIV vaccine

Angie Marsh\textsuperscript{1}, Aruna P. Ambagala\textsuperscript{2,3}, Justen Hoffman Russell\textsuperscript{1}, Jacqueline K. Chan\textsuperscript{3}, David O. Willer\textsuperscript{2,3}, Kelly S. MacDonald\textsuperscript{1,2,3}

\textsuperscript{1}Department of Immunology, University of Toronto, ON, Canada, M5S 1A8 \textsuperscript{2}Department of Medicine, University of Toronto, ON, Canada, M5S 1A8 \textsuperscript{3}Department of Microbiology, Mount Sinai Hospital, ON, Canada, M5G 1X5

Author Contributions: All experiments and figures described in this chapter were performed and generated by myself (A.K.M.). J.R.H. assisted me while performing the experiments for Section 4.5.5, and J.K.C. is the technician in the K.S.M. laboratory who helped to prepare and maintain reagents, resources, etc. D.O.W. generated the homologous recombination plasmid (described in section 4.5.2, and Figure 4-3A). A.K.M., A.P.A., D.O.W. and K.S.M. were involved with conceiving, designing, and analyzing the experiments.
4.1 Chapter Overview

To begin the development of CyCMV as a viral vector, the linear virus must be cloned into a circular plasmid. Once in a plasmid, the vaccine antigens can be incorporated into the virus and genetic manipulations or attenuations can be performed. However, due to the extremely large size of CyCMV, traditional cloning strategies cannot be employed. To circumvent this, I took advantage of a Bacterial Artificial Chromosome (BAC) to clone CyCMV as a circular plasmid, since this system can stably support the cloning of large DNA fragments. This chapter describes the generation of the required recombinant viruses for the various downstream cloning steps, and the successful cloning of CyCMV as a BAC.

Main Objective: To isolate, genetically characterize, and develop cynomolgus macaque cytomegalovirus (CyCMV) as an HIV/SIV viral vector for evaluation in a non-human primate model.

Specific Objectives

1. Isolation and characterization of cytomegalovirus from cynomolgus macaques (Chapter 2)

2. Identification and preliminary classification of the isolated virus as a novel CMV (Chapter 2)

3. Next-generation DNA sequencing of CyCMV (Chapter 3)

4. Bioinformatic assembly of the CyCMV DNA sequence reads (Chapter 3)

5. Genomic Analysis of CyCMV (Chapter 3)

6. Cloning of CyCMV as a Bacterial Artificial Chromosome
4.2 Abstract

Within the field of HIV vaccine design, it has been recognized that there is a real need to diversify the pool of vectors undergoing testing and to utilize vectors that can induce durable immunity. Herpesviruses, such as cytomegalovirus (CMV), show much promise as viral vectors as they establish lifelong latency with periodic reactivation. We hypothesize that reactivation of the CMV vector may enable persistent expression of vaccine antigens. To evaluate CMV-based HIV/SIV viral vectors in a non-human primate model, I have previously isolated and characterized a novel cynomologus macaque CMV (CyCMV). Like all herpesviruses, CyCMV has a large dsDNA viral genome of 218,041 bp that encodes 262 ORFs. In this regard, viral manipulation and the insertion of vaccine antigens cannot be performed using traditional cloning, however a Bacterial Artificial Chromosomes (BAC) can facilitate the cloning of large DNA inserts. Here I report the cloning of CyCMV as a BAC using in vitro Cre/loxP recombination. To mediate CyCMV BAC cloning using a Cre/loxP recombination system, loxP sites were inserted into CyCMV by homologous recombination using eGFP as a selection marker to generate CyCMV-eGFP. Following successful insertion of the loxP sites, the eGFP selection marker was excised by Cre/loxP recombination leaving in a single loxP site required for recombination with the BAC plasmid. The recombinant CyCMV was cloned as a BAC and plaque purified by selecting for eGFP encoded within the BAC plasmid to generate CyCMV-eGFP-BAC. Circularized CyCMV-eGFP-BAC was transformed into DH10B cells for the insertion of vaccine antigens and downstream modification of the viral genome using prokaryotic recombineering. In addition to our efforts of developing CyCMV a viral vector, CyCMV-eGFP may be utilized for visualization during CMV pathogenesis studies, and CyCMV-eGFP-BAC will enable the virus to be more accessible for CMV biologists allowing for viral manipulation studies, functional characterization of CMV genes, and HCMV vaccine development.
4.3 Introduction

Cytomegalovirus (HHV-5) is a human herpesvirus belonging to the Betaherpesvirus family. Cytomegalovirus (CMV) has a broad cellular and tissue tropism causing persistent and asymptomatic lifelong infection in immunocompetent host (reviewed in (553)). Although CMV characteristically causes a latent virus infection, it periodically reactivates throughout the life of the host and initiates a strong cellular immune response with greater than 10% of the total T cell repertoire being CMV-specific (232, 297, 810). The ability of CMV to reactivate and stimulate a robust immune response makes it an attractive viral vector to evaluate for an HIV vaccine. To this end, CMV-based HIV viral vectors are being evaluated in non-human primate (NHP) models prior to human clinical trials. Picker et al. have recently shown that using rhesus macaque CMV (RhCMV) as an SIV viral vector induces a strong effector memory immune response that is both durable and protective from SIV disease progression (334, 338) (reviewed in detail in Section 5.1.2 Rhesus Macaque Cytomegalovirus-Based Viral Vector). With these promising results, it is evident that now more than ever, a variety of CMV strains should be assessed in this context before moving forward to human clinical trials. In this regard, cynomolgus macaque CMV (CyCMV) has recently been isolated and genetically characterized in an effort to evaluate CMV-based vectors in an additional NHP model (22, 504). The development of CyCMV as a vaccine vector first requires circularizing the virus to facilitate the insertion of vaccine antigens, which can be mediated by cloning the virus as a Bacterial Artificial Chromosome (BAC).

Like all herpesviruses, CMVs are large linear double-stranded DNA viruses of >200kb making them difficult to manipulate in vitro. Given the immense size of CMV genomes, traditional cloning cannot be employed. To overcome this, BAC cloning can be performed in order to facilitate handling and manipulation of the virus. The BAC cloning system was first designed in 1992 (753) and since has been widely utilized to clone a number of large human herpesviruses including: Human Simplex virus-1 (HSV) (368, 719, 790) and HSV-2 (538), Varicella Zoster virus (VZV) (575), Epstein-Barr virus (EBV) (132, 210, 406), CMV (reviewed below), human herpesvirus-6A (HHV-6A) (84), and Kaposi’s Sarcoma-associated herpesvirus (KSHV) (211, 913), with human herpesvirus-7 (HHV-7) being the only herpesvirus that has yet to be cloned as a BAC. The very first herpesvirus to be cloned as a BAC was mouse CMV (MCMV) (6, 539) followed by a number of human CMV (HCMV) strains including: AD169 (87,
903), Towne (501), live attenuated Towne vaccine (183), Merlin (787), TB40/E (763), TR, PH, and FIX (VR1814) (572). Among other species, rhesus macaque CMV (RhCMV) (128) and guinea pig CMV (GPCMV) (526) have also been cloned as a BAC.

The BAC cloning system utilizes prokaryotic genetics through an F-plasmid, which encodes the fertility factor (F factor) that restricts only one or two BAC plasmids to be introduced into each bacterium, thus limiting unwanted DNA recombination (753). BAC plasmids are stably maintained in *Escherichia coli* (*E. coli*) and are capable of inserting DNA fragments of greater than 300kb in length (753). CMV can be cloned into a BAC using an *in vitro* Cre/loxP recombination system (793). However, given the large size of HCMV, it can only accommodate the insertion of ~5 kb of exogenous DNA into the viral genome in order to effectively encapsulate the virion during replication in mammalian cells (539, 903). A BAC plasmid can range from 7-9 kb in length and thus *in vitro* manipulation of the virus would likely result in the spontaneous deletion of viral genes and/or the BAC plasmid. The benefit of using a BAC cloning system is that all mutagenesis of the circularized virus can be done by homologous recombination in *E. coli*, allowing for a more stable, rapid, and efficient manipulation of the virus. Upon completion, the BAC plasmid can be transfected into mammalian cells enabling full reconstitution of infectious virus (621).

Next-generation sequencing and characterization of the CyCMV genome (discussed in Chapter 3) has facilitated the identification of target sites for BAC cloning. Here I describe the development of a recombinant eGFP-expressing CyCMV (CyCMV-eGFP) and the cloning of CyCMV as a bacterial artificial chromosome (CyCMV-eGFP-BAC). With the ultimate goal of developing CyCMV as a viral vector for an HIV/SIV vaccine, the cloning of CyCMV as a BAC will facilitate the insertion of vaccine antigens as well as any viral manipulations required to attenuate the vector. In addition to our efforts of developing CyCMV a viral vector, CyCMV-eGFP-BAC will enable the virus to be more accessible for CMV biologists allowing for virus manipulation studies (insertions, deletions, mutations), functional characterization of CMV genes, and HCMV vaccine development.
4.4 Materials and Methods

4.4.1 Cells and Virus

CyCMV Mauritius strain was isolated from a 15-year-old cynomolgus macaque of Mauritius origin (described in Chapter 2) and CyCMV Ottawa strain was isolated from a 4-year-old cynomolgus macaque of Filipino origin (described in (22) and Chapter 3). The viruses were propagated and titred on telomerase-immortalized rhesus macaque fibroblast cells (Telo-RF), generously provided by Drs. Peter Barry and William Chang (University of California, Davis) (432), and maintained in Dulbecco’s Modified Eagle’s Medium (DMEM; Sigma) supplemented with 10% FBS (Wisent Bioproducts), 100 U/ml penicillin, and 100 µg/ml streptomycin (Sigma). Human umbilical vein endothelial cells (HUVEC), kindly provided by Dr. Kevin Kain (University of Toronto, Canada), were maintained in Endothelial cell Growth Media (EGM-2) MV BulletKit (Lonza) supplemented with 10% FBS, 100 U/ml penicillin, and 100 µg/ml streptomycin.

4.4.2 Generation of CyCMV-eGFP Recombinant Virus

CyCMV (Ottawa strain) cyUS1 and cyUS2 (504) were cloned into a pWC132 plasmid encoding an eGFP cassette flanked by two loxP sites (generously provided by Drs. W.L. William Chang and Peter Barry) (Figure 4-3). Telo-RF cells were seeded in a 6-well tissue culture plate and transfected with 3.2 µg of plasmid DNA using Lipofectamine 2000 Transfection Reagent (Invitrogen), as per manufacture’s protocol. Twenty-four hours post-transfection, the cells were infected in duplicates with CyCMV (Ottawa or Mauritius strain) at a multiplicity of infection (MOI) of 0.2, 2, or 20, and incubated at 37°C in a 5% CO₂ incubator. Once 100% CPE was observed, the cells and supernatant were harvested and frozen at -80°C with 0.2 M sucrose. Viral supernatant was used to infect Telo-RF cells in a 6-well plate using 10-fold serial dilutions and the plate was incubated at 37°C for 2 hrs. The inoculum was removed and the cells were overlaid with a 1:1 mixture of 2X MEM (containing 10% FBS and 4 mM L-glutamine; Quality Biologicals Inc.) and 2% low-melting agarose (type VII; Sigma). Individual eGFP-positive plaques were selected, 6 rounds of plaque purification were performed, and the purified recombinant CyCMV-eGFP recombinant virus was propagated on Telo-RF cells.
4.4.3 Cre/loxP Recombination Methodology

The Cre/loxP recombination system is extensively utilized in our \textit{in vitro} BAC cloning process, and thus warrants a brief explanation of the methodology. LoxP is a 34 bp sequence with two 13 bp inverted repeat sequences flanking an internal 8 bp spacer sequence (Figure 4-1A). Cre recombinase protein (38 kDa) is derived from bacteriophage P1, and acts as a catalyst for loxP recombination by binding to the recognition sites and cleaving loxP at defined sites in the spacer sequence (Figure 4-1A) (793). Depending on the orientation of the loxP sites, Cre recombination can facilitate gene deletion, translocation, or inversion. For our purposes, I will be using Cre/loxP recombination to remove eGFP and for recombination with the BAC plasmid, an example of each application is shown in Figure 4-1.

\textit{Figure 4-1 Cre/loxP Recombination System.} The loxP sites (shown in red) serve as recognition sequences for the Cre recombinase enzyme to bind and cleave the DNA, at the designated cut sites located within the spacer sequence (A). This cleavage can facilitate gene deletion, leaving a single loxP site (A), or mediate recombination between two loxP sites for plasmid insertion (B).
4.4.4 Generation of CyCMV-loxP

To facilitate Cre-loxP recombination (Figure 4-1), Telo-RF cells were seeded in a 6-well plate and transfected with a Cre-expressing plasmid pWC205 (Figure 4-3) (kindly provided by Drs. W.L. William Chang and Peter Barry) at 3.2 µg/well using Lipofectamine 2000 Transfection Reagent, as per manufacturer’s protocol. Twenty-four hours post-transfection, the cells were infected with 10-fold serial dilutions of CyCMV-eGFP from clarified supernatant. Once 100% CPE was observed, the cells and supernatant were harvested and frozen at -80°C with 0.2 M sucrose. Viral supernatant was used to infect Telo-RF cells in a 6-well plate using 10-fold serial dilutions and the plate was incubated at 37°C for 2 hrs. The inoculum was removed and the cells were overlaid with a 1:1 mixture of 2X MEM (containing 10% FBS and 4 mM L-glutamine) and 2% low-melting agarose (type VII). Individual eGFP-negative plaques were selected and propagated on Telo-RF cells for 4 rounds of plaque purification. The purified CyCMV-loxP recombinant virus was propagated, concentrated on a sorbitol cushion (20% D-Sorbitol, 50 mM Tris-Cl pH 7.2, 1 mM MgCl₂) by centrifugation at 71,000 x g for 1.5 hrs at room temperature, and titred on Telo-RF cells using a standard plaque assay.

4.4.5 Cloning CyCMV as a Bacterial Artificial Chromosome

Telo-RF cells were seeded in a 6-well plate and co-transfected with 3.2 µg total/well at a 3:2 ratio with pWC205 and the BAC plasmid pWC155 (Figure 4-3) (generously provided by Drs. W.L. William Chang and Peter Barry), respectively, using Lipofectamine 2000 Transfection Reagent, as per manufacturer’s protocol. Twenty-four hours post-transfection, the co-transfected cells were infected with 0.3 or 0.03 MOI of concentrated CyCMV-loxP. At 48 or 96 hrs post-infection, the cells and supernatant were harvested and frozen at -80°C with 0.2 M sucrose. Viral supernatant was used to infect Telo-RF cells in a 6-well plate using 10-fold serial dilutions and the plate was incubated at 37°C for 2 hrs. The inoculum was removed and the cells were overlaid with a 1:1 mixture of 2X MEM (containing 10% FBS and 4 mM L-glutamine) and 2% low-melting agarose (type VII). Individual eGFP-positive plaques were selected and propagated on Telo-RF cells for 4 rounds of plaque purification. The purified CyCMV-eGFP-BAC recombinant virus was propagated, concentrated on a sorbitol cushion (20% D-Sorbitol, 50 mM Tris-Cl pH 7.2, 1 mM MgCl₂) by centrifugation at 71,000 x g for 1.5 hrs at room temperature, and titred on Telo-RF cells using a standard plaque assay.
4.4.6 Viral DNA Isolation

Viral DNA isolation for the CyCMV-eGFP and CyCMV-loxP recombinant viruses was performed as described previously (22, 504). CyCMV-eGFP-BAC viral DNA was isolated with a similar methodology (621). Briefly, infected cells were centrifuged at 200 x g for 5 mins at room temperature, the cell pellet was resuspended in Hirt extraction buffer (10 mM Tris-HCl, 10 mM EDTA, pH 8.0) (361), centrifuged again, and the pellet was lysed with Hirt extraction buffer and 250 µg/ml proteinase K (Qiagen). Sodium dodecyl sulfate (0.6%) and 1 M sodium chloride were added to the lysate and incubated at 4°C overnight. The lysate was clarified by centrifugation at 9,300 x g for 30 mins at 4°C and proteins were removed from the supernatant with two phenol (Sigma) and two chloroform:isoamyl alcohol (24:1; Sigma) extractions. Viral DNA was precipitated with 2 volumes of ethanol, centrifuged at 2,300 x g for 2 mins at room temperature, washed once with 70% ethanol, and dissolved in 10 mM Tris-HCl (pH 8.0).

4.4.7 Electroporation

CyCMV-eGFP-BAC isolated viral DNA was transformed into MegaX DH10B electrocompetent E. coli cells (Invitrogen) by electroporation. Electroporation was performed using 1 mm electroporation cuvettes (Bio-Rad) and a Bio-Rad Gene Pulser under the following settings: voltage 1.8 kV, resistance 200 OHMS, capacitance 25 µF, and 4.4 ms time constant. Recovered cells were incubated for 1 hr at 37°C in 1 ml of recovery medium (Invitrogen), centrifuged at 20,000 x g for 5 mins at room temperature, and resuspended in 80 µl of recovery medium. The sample was plated out onto agar plates containing 15 µg/ml of chloramphenicol (CAM) and incubated at 37°C until CAM-resistant colonies were observed.

4.4.8 Polymerase Chain Reaction

To confirm the insertion of eGFP and loxP between cyUS1 and cyUS2, oligonucleotide primers were designed within cyUS1 and cyUS2 to amplify the intergenic region, forward primer 5’-ACCGGACTTTTTGCGCATGACT-3’, reverse primer 5’-GCCAGTCCCCGCACTGAC-3’. To confirm the insertion of the BAC plasmid, the following primers were designed to target genes encoded in the pWC155 plasmid, including the chloramphenicol resistance gene F: 5’- CCCGCCCTGCCACTCATCG-3’, R: 5’-ACATTTTGAGGCATTTCAGTGCAGTTGC-3’, and a region spanning a portion of both repE and parA F: 5’-CTGACCTCCGCCGCCGCTTC-3’, R: 5’-GGGCAGCAACCCCGATCACC-3’, or in a region of DNA polymerase located with
CyCMV F: 5'-TTGTCAAGAACACCGTGCGCAAC-3', R: 5'-AATGAGAGCTGCGGCGGACA-3'. PCR conditions were as follows: 95°C for 10 mins, 30 cycles of 95°C for 30 sec, 55°C for 30 sec, 72°C for 3 mins, followed by a 7 mins extension at 72°C. PCR amplicons were cloned into pCR-Blunt II-TOPO vector (Invitrogen) and transformed into TOP10 chemically competent cells (Invitrogen). The full-length inserts were sequenced by Sanger sequencing using standard vector primers (M13 For (UP): 5'-CACGACGTTGTAAAACGAC-3' and M13 Rev (-27) - Invitrogen: 5'-CAGGAAACAGCTATGAC-3') and the sequences was aligned and analyzed using Geneious Pro v.5.6.

4.5 Results

4.5.1 CyCMV-BAC Cloning Process

The genomic sequencing and characterization of CyCMV (discussed in Chapter 3) enabled the identification of optimal locations for targeting the insertion of the BAC. In view of not disrupting gene transcription and translation, the 235 bp intergenic region between two tail-to-head open reading frames, cyUS1 and cyUS2 was chosen as the BAC cloning site. Furthermore, previous studies have shown that recombination in the cyUS1-cyUS2 homologous region of the RhCMV genome did not impede viral replication or pathogenesis (128, 130). A multi-step in vitro cloning procedure was employed to clone CyCMV as a BAC with a focus on maintaining the integrity of the viral genome throughout the process. The cloning took advantage of previously developed in vitro recombination systems however, given that in vitro recombination is a rare event, an enhanced green fluorescent protein (eGFP) was used as a selection marker to visualize successful recombination, by selecting for or against, throughout the cloning process (Figure 4-2).
Figure 4-2 Schematic Representation of Multi-Step CyCMV-BAC Cloning Process. A) Two loxP sites were inserted into the CyCMV genome by homologous recombination with pWC132, using eGFP as a selection marker, to generate CyCMV-eGFP. B) pWC205 expressing Cre recombinase was used to facilitate Cre/loxP recombination resulting in the excision of eGFP and the generation of a single loxP site to create CyCMV-loxP. C) The BAC plasmid (pWC155) encoding an eGFP selection marker was inserted into the viral genome via Cre/loxP recombination generating CyCMV-eGFP-BAC.

To facilitate the downstream cloning of CyCMV as a BAC using Cre/loxP recombination (described in Section 4.4.3), a loxP recognition sequence was incorporated into viral genome. To this end, two recombinant viral intermediates were generated, CyCMV-eGFP and CyCMV-loxP. To create CyCMV-eGFP, an eGFP marker flanked by two loxP sites was inserted into the viral genome by homologous recombination (Figure 4-2A). Subsequently, to generate a single loxP site for BAC recombination, the eGFP marker was excised and the loxP sites were recombined by Cre/loxP recombination to form a single recognition sequence in CyCMV-loxP (Figure 4-2B).
The BAC plasmid encoding an eGFP selection marker was cloned into the viral genome by Cre/loxP recombination to successfully generate a CyCMV-eGFP-BAC recombinant virus in which the circularized virus can be transferred into a bacterial system for viral manipulation and HIV/SIV vaccine development (Figure 4-2C).

4.5.2 Recombination Plasmids

The recombination plasmids used throughout the cloning process and the BAC cloning vector are outlined in Figure 4-3. To maximize mRNA expression and stability, the genes required for selection and recombination were driven by an SV40 early promoter and terminated by a SV40 early mRNA polyadenylation signal. To construct the homologous recombination plasmid, cyUS1 and cyUS2 were PCR amplified from CyCMV Ottawa strain (504), sequenced to ensure fidelity of amplicon, and subcloned into the pWC132 plasmid flanking the loxP-eGFP-loxP cassette (Figure 4-3A). The inserted cyUS1 homologous region was 797 bp in length, which spanned the complete cyUS1 ORF (507 bp) and the surrounding areas. Similarly, the cyUS2 region covered the sequence upstream of cyUS2 and the majority of the cyUS2 ORF (501 bp) to comprise 620 bp of homologous sequence. We proposed that these homologous regions would be sufficient to achieve recombination since low-level recombination has been observed in mammalian cells with as little as 14 bp of homologous sequence (707). Following the insertion of the loxP sites, Cre/loxP was used to facilitate recombination with a Cre recombinase expressing plasmid (pWC205) (Figure 4-3B) (128). Furthermore, the pWC205 plasmid encodes a synthetic intron, which prevents functional Cre protein expression in E. Coli (768). The BAC plasmid (pWC155) encodes an F-factor replication origin (redF) and the regulatory genes (parA, parB, parC, repE) required for stable maintenance and single-copy replication of the BAC DNA in E. Coli (753, 863). For recombination purposes, pWC155 also encodes a single loxP site and eGFP marker for in vitro recombination and selection, as well as a chloramphenicol resistance gene for prokaryotic selection of successfully cloned viral genomes (Figure 4-3B) (128).
Figure 4-3 CyCMV-BAC Cloning Plasmids. A) To facilitate homologous recombination with parental CyCMV, cyUS1 and cyUS2 were subcloned into the pWC132 plasmid encoding eGFP flanked by two loxP sites. B) The Cre-expressing plasmid pWC205 encodes a Cre recombinase enzyme that will serve to facilitate Cre/loxP recombination. The BAC plasmid pWC155 encodes a single loxP site, eGFP as an in vitro selection marker, and a chloramphenicol resistance gene for prokaryotic selection.

4.5.3 Generation of a CyCMV-expressing eGFP Recombinant Virus

Both strains of CyCMV (Ottawa and Mauritius) were utilized in parallel for the generation of CyCMV-eGFP. Telo-RF cells were transfected with 3.2 µg of pWC132 and twenty-four hours post-transfection, the cells were infected in duplicates with CyCMV at an MOI of 0.2, 2, of 20. Once 100% CPE was observed, the cells and supernatant were harvested and viral supernatant was used to infect Telo-RF cells using 10-fold serial dilutions. Individual eGFP-positive plaques could be observed for both CyCMV strains, however CyCMV Ottawa displayed fewer plaques and the propagation of the recombinant Ottawa strain virus could not be maintained during plaque purification. As a result, all downstream cloning was performed using CyCMV Mauritius strain, in which individual eGFP-positive plaques were successfully selected (Figure 4-4A) and 6 rounds of plaque purification were performed. Purified CyCMV-eGFP recombinant virus was propagated on Telo-RF cells and eGFP expression was uniformly
observed across the infected monolayer (Figure 4-4B). To confirm the recombinant virus retained the homologous CyCMV genes shown to be required for productive infection of endothelial cells (cyUL128, cyUL130, cyUL131A), HUVECs were infected with CyCMV-eGFP. Productive infection of HUVECs was observed, as demonstrated by eGFP-positive CyCMV cytopathic effect (Figure 4-4C), suggesting that the insertion of the loxP-eGFP-loxP cassette and subsequent in vitro propagation did not interfere with the CyCMV encoded cellular tropism genes, however this will require confirmation by next-generation sequencing of the viral genome.

![Figure 4-4 CyCMV-Infected Cells Expressing eGFP.](image)

CyCMV-eGFP recombinant virus was inoculated on confluent monolayers and exhibited CyCMV cytopathic effect expressing eGFP. A) Higher magnification of CyCMV-eGFP-infected Telo-RF cells demonstrating a CyCMV-eGFP plaque selected for during plaque purification. Following plaque purification, purified CyCMV-eGFP recombinant virus was inoculated on Telo-RF (B) or HUVEC (C) cell monolayers and exhibited widespread cytopathic effect.

The fidelity of the inserted loxP-eGFP-loxP cassette was confirmed by PCR amplification and sequencing of the intergenic region between cyUS1 and cyUS2. Sequencing confirmed the integrity of the inserted cassette, however a newly generated G to A mutation was discovered in the loxP site located between eGFP and cyUS2 (Figure 4-5). The mutation occurred at position 6 in 3’ inverted recognition site of loxP. Previous studies have demonstrated that the base pair located at position 6 was the most sensitive for Cre recombinase binding to the recognition site and symmetrical mutations at these position greatly reduce recombination efficiency (341), while other studies have shown that targeted mutagenesis at position 6 did not impeded Cre/loxP recombination (727). Based on these conflicting reports, I proceeded with the downstream cloning and elucidation of the impact of the loxP mutation on recombination. All told, the
generation, purification, and propagation of this recombinant virus took nine months before the virus was ready to move to the next phase of the cloning process.

**Figure 4-5 Mutation in LoxP Recognition Site.** PCR amplification and sequencing of the inserted loxP-eGFP-loxP region within CyCMV-eGFP revealed a G to A mutation at position 6 in the recognition sequence of the loxP site located between eGFP and cyUS2.

### 4.5.4 Generation of CyCMV-loxP Intermediate

In order to facilitate downstream insertion of the BAC plasmid, a single loxP site in the BAC-targeting region is required, thus an intermediate CyCMV-loxP recombinant virus was generated (Figure 4-2B). To produce Cre-expressing fibroblast cells, Telo-RF cells were transfected with 3.2 µg/well of the Cre-expressing plasmid (pWC205). Twenty-four hours post-transfection, the cells were infected with 10-fold serial dilutions of CyCMV-eGFP to facilitate Cre/loxP recombination. After 6 days of infection, 100% CPE was observed at the highest dilution, the virus was harvested, and clarified viral supernatant was used for plaque purification. Though the mutation in the loxP recognition site of the CyCMV-eGFP virus may have impeded the efficiency of Cre/loxP recombination, CyCMV eGFP-negative plaques were observed signifying that successful recombination had occurred and eGFP was successfully excised. Throughout the six-month process, four rounds of plaque purification were performed with selection for CyCMV-loxP plaques that were not expressing eGFP and the purified virus was propagated on Telo-RF cells. To confirm the insertion of a single 34 bp loxP site between cyUS1 and cyUS2, sequence analysis was performed and no mutations in the loxP sequence were present at this juncture, suggesting that the previously observed G to A mutation was removed during recombination of the two loxP sites (Figure 4-6). Following viral concentration and titration, CyCMV-loxP achieved high viral titers (7.8x10^7 PFU/ml), which were comparable to
wild type CyCMV titres suggesting that the insertion of the single loxP site did not impede viral replication.

Figure 4-6 CyCMV-loxP DNA Sequence. PCR amplification and sequence analysis confirmed the presence of a single 34 bp loxP site inserted between cyUS1 and cyUS2 of CyCMV-loxP. The sequence was aligned and annotated using Geneious Pro v. 5.6 (227).

4.5.5 Cloning CyCMV as a Bacterial Artificial Chromosome

The single loxP site located within the CyCMV-loxP recombinant virus facilitated BAC cloning using Cre/loxP recombination with the pWC155 plasmid (Figure 4-2C). Telo-RF cells were co-transfected with 3.2 µg total/well at a 3:2 ratio with Cre-expressing pWC205 plasmid and the BAC plasmid pWC155, respectively (Figure 4-3B). Twenty-four hours post-transfection, the co-transfected cells were infected with 0.6 or 0.06 MOI of concentrated CyCMV-loxP. At 48 or 96 hrs post-infection, the cells and supernatant were harvested and viral supernatant was used for plaque purification. Evidence that CyCMV-loxP successfully recombined with the BAC plasmid to generate a CyCMV-eGFP-BAC recombinant virus was visualized by the presence of eGFP-positive plaques 5 days following limiting dilution infections. There was no observed difference between MOIs used for infection or the duration of infection prior to harvest. eGFP-positive plaques were selected and propagated on Telo-RF cells for 4 rounds of plaque purification in an effort to minimize in vitro propagation. All told, fourteen clones were generated and plaque purified to different levels in the event that the insertion of the BAC spontaneously produced CyCMV gene deletions, which will not be determined until next-generation genomic sequencing is performed. After eight months of plaque purification and propagation, purified CyCMV-eGFP-BAC was concentrated and titred on Telo-RF cells using a standard plaque assay.
Following successful cloning of CyCVM as a BAC, the recombinant viral genome was transferred into a bacterial system for downstream viral manipulations and vaccine development. To mediate the transformation of CyCMV-eGFP-BAC into *E. Coli*, Telo-RF cells were infected with 0.06 MOI of CyCMV-eGFP-BAC for 120 hrs in an effort to maximize the number of circularized viral genomes undergoing rolling circle replication. Circularized virus was isolated by Hirt extraction and purified viral DNA was transformed into DH10B *E. coli* cells by electroporation (361). Positive transformants were selected on chloramphenicol *E. coli* cells and screened by colony PCR (Figure 4-7). To confirm that the cloning process did not cause any gross changes to the viral genome, the CyCMV-eGFP-BAC plasmid was restriction enzyme digested (Figure 4-8). The digestion confirmed the integrity of the CyCMV genome and demonstrated an intact viral genome with a similar band pattern to the previously digested parental strain (Figure 2-2). In addition to this gross analysis, next-generation sequencing of CyCMV-eGFP-BAC will provide a more detailed genomic examination to confirm any genetic mutations or gene rearrangements.
Figure 4-7 Screening CyCMV-eGFP-BAC Colonies. Primers specific for genes encoded in the pWC155 BAC plasmid (refer to Figure 4-3) including A) the chloramphenicol resistance gene (amplicon: 595 bp) and B) a region spanning a portion of both repE and parA (amplicon: 1,079 bp) were used to screen positive clones. C) Primers designed within CyCMV DNA polymerase (amplicon: 416bp) were used to confirm the insertion of CyCMV in positive BAC colonies. The pWC155 BAC plasmid (A and B), and the parental strain of CyCMV (C) were used as positive controls. Lane MW, 1 kb DNA ladder.

Figure 4-8 Restriction Enzyme Digestion of CyCMV-eGFP-BAC. CyCMV-eGFP-BAC was restriction enzyme digested with BamHI, HindIII, and EcoRI, and fractionated on a 0.8% agarose gel. Lane MW, 1 kb DNA ladder.

4.6 Discussion

In addition to Bacterial Artificial Chromosomes (BACs), a number of different cloning systems that have been developed to clone large DNA fragments, including Yeast Artificial Chromosomes (YACs) (107), P1-derived Artificial Chromosomes (PACs) (379), Mammalian Artificial Chromosomes (MACs) (339, 378), and cosmids (162, 380), however these methods
present limitations with cloning large herpesviruses. YACs are capable of cloning extremely large DNA fragments of up to 1000 kb however, they are maintained in yeast and are prone to spontaneous rearrangements and the generation of chimeric clones during propagation (107, 322, 557). Similar to BACs, cosmids and PACs have the advantage of being bacterial-based cloning vectors, although their utility for cloning large herpesviruses is limited as they can only insert 35-45 kb and 100-300 kb, respectively (379, 380). Given the unstable nature of the repetitive sequences in large herpesvirus genomes, BACs represent the superior method of cloning as they stably maintain up to 300 kb DNA fragments in a bacterial host and encode an F-factor gene that serves to maintain a single plasmid copy per bacterial cell, thus limiting spontaneous viral DNA recombination during propagation (753).

To facilitate the cloning of CyCMV as a BAC by Cre/loxP recombination, a CyCMV-eGFP recombinant virus was generated. In addition to the utility of CyCMV-eGFP as an intermediate virus required in the BAC cloning process, this novel recombinant virus represents a visualization tool that may be used for in vivo studies examining CMV biology and pathogenesis in a cynomolgus macaque model. Initially, the in vivo replication kinetics and latency properties of this recombinant virus will need to be experimentally evaluated in cynomolgus macaques. The eGFP-expressing virus will aid in visualizing infection with implications for evaluating CyCMV-specific cellular and tissue tropism, viral shedding, reactivation, and localization. Similar studies in rhesus macaques have documented that the insertion of eGFP into RhCMV (RhCMV-eGFP) did not impede viral pathogenesis and furthermore, RhCMV-eGFP allowed for visualization of CMV infection patterns in the brain of fetal rhesus macaques (130). Given that CyCMV is a novel virus, CyCMV-eGFP will facilitate the first efforts to study of CyCMV pathogenesis in cynomolgus macaques, which may provide valuable information regarding host-pathogen interactions that may direct future evaluations of CMV-based HIV/SIV vaccines in this model.

There are many stages during the process of cloning large herpesvirus genomes as bacterial artificial chromosomes in which the virus may delete genes that are not critical for viral fitness. Given the large size of CMV genomes, they can only accommodate the insertion of ~5kb of exogenous DNA into the viral genome in order to effectively encapsulate the virion during replication in mammalian cells (539, 903). The BAC plasmid (pWC155) cloned into CyCMV was 8,807 bp in length and thus in vitro propagation of CyCMV-eGFP-BAC may have resulted
in the spontaneous deletion of viral genes that are not critical for viral fitness. In this regard, next-generation sequencing of the CyCMV-eGFP-BAC viral genome will need to be performed to determine if any gene deletions or rearrangements have occurred, and all open reading frames (ORFs) will require annotation. A thorough comparison of the genomic organization between the CyCMV parental strain and BAC cloned viral genome will determine if the cloning process resulted in any gene rearrangements, mutations, gene deletions, as well as other tissue culture-adapted changes. In the event that genes were deleted, the functional impact of the deletion will need to be elucidated. It is unlikely that the genes involved in viral replication would be compromised, as this would impede the viral fitness of the virus and comparable wild-type CyCMV titres were observed throughout the cloning process.

Given that the all stages of the in vitro cloning process were performed in fibroblast cell lines, it is possible that viral tropism genes would be removed, such as HCMV homologues (UL128, UL130 and UL131A) that form complexes with gH and gL to facilitate infection of endothelial cell, macrophage and dendritic cells (332, 711, 712, 762). Previous analysis of BAC cloned HCMV and RhCMV genomes have demonstrated limited genetic variability in the BAC-derived viruses compared to the parental strains (494, 787). In both the HCMV- and RhCMV-BAC genomes, 12 and 13 nucleotide mutations were observed, respectively, with the majority being tissue culture adapted mutations located in RL13, UL36, and UL128 (196, 494, 787). CyCMV does not encode a RL13 homologue (Table 3-2), however it does encode homologues for the anti-apoptotic gene UL36, and the cellular tropism gene UL128 (Table 3-1), thus it will be particularly important to pay specific attention to those genes when analyzing the CyCMV-eGFP-BAC genome.

Prior to proceeding with the insertion of the vaccine antigens and the generation of the vaccine constructs, further assessment of CyCMV-eGFP-BAC will be required. To confirm that the CyCMV-eGFP-BAC plasmid has retained infectiousness in vitro and to evaluate the impact of the insertion of the BAC sequence on viral growth kinetics, the CyCMV-eGFP-BAC plasmid must be transfected into mammalian fibroblast cells. Upon successful transfection, the plasmid should generate infectious reconstituted CyCMV-eGFP-BAC (rCyCMV-eGFP-BAC) progeny virus. With respect to viral growth kinetics, it is essential to evaluate the effect of the inserted BAC sequence on the ability of the virus to replicate in vitro. A comparison between the parental CyCMV, the intermediate cloning viruses (CyCMV-eGFP, CyCMV-loxP), and rCyCMV-eGFP-
BAC will determine if the insertion of the BAC had a negative impact on viral growth kinetics. Previous evaluations of growth kinetics with reconstituted HCMV-BAC (AD169 strain) and RhCMV-BAC (68.1 strain) demonstrated a notable reduction in viral titres in comparison to wild type (128, 903), whereas HCMV-BAC (Towne strain) did not differ from wild type HCMV (233). However, when the BAC sequence was excised from the strains that did show impairment, the growth kinetics mirrored that of wild type CMV (128, 903). This is what would be expected given that CMVs can only accommodate the insertion of ~5kb of additional DNA in the viral genome without disrupting encapsidation, however this will not be problematic given that the vaccine vectors will have the BAC sequence excised following the insertion of the vaccine antigens and prior to vaccination (discussed in Section 5.2.2.2 Generation of Self-Excisable BAC). CyCMV-eGFP-BAC was constructed to allow the virus to be manipulated into a self-excisable BAC as two loxP sites flank the BAC plasmid, thus transfection into Cre-expressing cells will remove the BAC and result in reconstituted CyCMV (CyCMVr) with only 34 bp of nonviral sequence (loxP site) remaining. Thus, for the downstream application of CyCMV as a viral vector for an HIV/SIV vaccine, the potentially reduced viral growth kinetics in the BAC cloned virus will likely not affect the final BAC-excised CyCMV vaccine constructs, although this will need to be experimentally determined. In support of this, in vivo studies using BAC-derived RhCMV have shown that the reconstituted RhCMV is infectious, replication competent, and pathogenic in rhesus macaques (128, 334, 338).

The evaluations of herpesviruses, such as varicella zoster virus and cytomegalovirus, as viral vectors for an HIV vaccine have shown much promise in non-human primate models. In particular, recent studies evaluating RhCMV-based viral vectors in vaccinated rhesus macaque have shown that the vaccine induced a strong effector memory immune response that was both durable and protective from SIV disease progression (334, 338) (reviewed in Section 5.1.2 Rhesus Macaque Cytomegalovirus-Based Viral Vector). In an effort to advance the development and evaluation of safe and effective HIV vaccine candidates, it has become evident that now more than ever CMV and a variety of CMV strains should be assessed in non-human primate models before moving forward to human clinical trials. The strong species-specificity of CMV highlights the importance of emerging non-human primate models to examine CMV-based HIV/SIV vaccines to provide additional models to compare and complement safety and efficacy outcomes prior to initiating human clinical trials. In this regard, I have cloned CyCMV as a BAC
to facilitate the insertion of the vaccine antigens and the manipulation of the viral genome for its development as an HIV/SIV vaccine vector. In addition to its utility for an HIV/SIV vaccine, CyCMV-eGFP-BAC will facilitate the handling and accessibility of the virus for CMV researchers to perform gene manipulation studies, evaluate CMV pathogenesis, immunogenicity, and vaccine design.

4.7 Acknowledgments

We thank Drs. Peter Barry and William Chang (University of California, Davis) for generously providing the cloning plasmids and the Telo-RF cells, and Dr. Kevin Kain (University of Toronto) for the HUVEC cells. We also thank Dr. William Chang, Dr. Dong Yu, and Dr. Kening Wang their technical assistance. A.K.M. was supported by The Queen Elizabeth II/Community Health Graduate Scholarships in Science and Technology, and the Dr. Bernhard Cinader Graduate Scholarship in Immunology. D.O.W. was supported by a Junior Investigator Development Award from the Ontario HIV Treatment Network (OHTN). A.A. and K.S.M. also received funding from the OHTN as a postdoctoral fellowship and career scientist award, respectively. This research was funded in part by the Canadian Institutes of Health Research (CIHR).
Chapter 5

5 Discussion and Future Directions
5.1 Discussion

5.1.1 The Development of CyCMV as an HIV/SIV Vaccine Viral Vector

Despite major advances in antiretroviral therapy and novel research focused on therapeutic methods for eradicating HIV, the HIV/AIDS pandemic continues to spread. An HIV vaccine remains the best method for controlling the HIV pandemic and it is imperative that vaccine research be a global priority. A successful HIV vaccine must be able to stop HIV replication as early as possible to inhibit HIV at the portal of entry and within the window of vulnerability, before the virus leaves the site of primary infection and undergoes a peak burst of viremia leading to peripheral dissemination. In this regard, we propose that a vaccine designed to establish a persistent anti-HIV effector memory T cell response at the mucosa will impede HIV in its earliest stage, and we believe that a herpesvirus-vectored vaccine, such as cytomegalovirus, has the ability to mount this type of immune response. Cytomegalovirus establishes life-long infection in the host, with evidence of repeated subclinical reactivation and immunogenicity, which could theoretically facilitate episodic expression of vaccine antigens. This ability of the virus to “self-boost” represents a significant benefit over other HIV vaccine vectors currently in clinical trials and may help address the key issue of transient and waning vaccine immunity. Thus, we hypothesize that using a chronic reactivating CMV, as an HIV vaccine vector will provide durable and persistent in situ mucosal T cell effector memory responses by eliciting protective effectors at the portal of HIV entry and in a timeframe most appropriate to inducing clinically relevant protection. In addition, we predict that we will be able to evaluate the utility and effectiveness of CMV as a viral vector in a cynomolgus macaque non-human primate/SIV challenge model.

To being the evaluation of a CMV-based HIV viral vector in the cynomolgus macaque model, the previously characterized rhesus macaque cytomegalovirus (RhCMV) was tested for cross-species infectivity in our cynomologus macaque model. Our results suggest that it may be inherently difficult to cross the species-specific barrier and infect cynomolgus macaques with RhCMV and that species-specificity could not be overcome between these closely related macaques. Thus, in order to evaluate CMV-based HIV vaccines in a cynomolgus macaque model, host-specific CMV was required. To this end, we isolated and characterized a novel
cynomolgus macaque cytomegalovirus (CyCMV) derived from urine co-cultures of macaques from Filipino (Ottawa strain) or Mauritius origin (Mauritius strain). CyCMV shares a similar size and morphology to previously characterized cytomegalovirus from other species, as observed by electron microscopy of virally infected cells. Furthermore, gross viral genome structure was comparable to RhCMV, and like all herpesviruses, CyCMV productively infects, replicates, and downregulates MHC class I expression in fibroblast cell lines derived from humans, rhesus macaques, and cynomolgus macaques. Phylogenetic analysis was performed to determine the phylogenetic relationship between CyCMV and other closely related CMVs. It was observed that CyCMV clustered with Old World Primates, specifically RhCMV, however the protein sequence of CyCMV glycoprotein B exhibits some unusual heterogeneity. In this locus, CyCMV Mauritius strain is more closely related to both strains of RhCMV than to CyCMV Ottawa strain, which is uncommon given that glycoprotein B is typically highly conserved. We hypothesize that the observed differences in this region between the two strains of CyCMV may be attributed to immune selective pressure on the virus following the geographic separation of these two populations of cynomolgus macaques. Complete genomic characterization of the Mauritius strain will aid in elucidating whether this divergence was unique to glycoprotein B, or if it is mirrored throughout other highly immunogenic genes encoded in the viral genome.

I performed next-generation sequencing of the CyCMV viral DNA isolated from CyCMV Ottawa and Mauritius strains. I comprehensively characterized the viral genome of CyCMV (Ottawa strain), which is 218,041 bp in length and encodes 262 open reading frames, 39 of which are homologues of the 40 core genes encoded by all herpesviruses (Figure 3-3) (553). Consistent with other cytomegalovirus genomes, CyCMV is organized with a unique long region followed by a unique short region (Figure 1-12). When comparing the genomic sequence and gene products of CyCMV to other cytomegaloviruses, CyCMV is largely collinear with the two published genomes of RhCMV. At the nucleotide level, the CyCMV genome is 54.8% identical to HCMV AD169 strain (131), 53.6% to chimpanzee CMV (198), 89.8% to RhCMV 68.1 strain (337), and 88.2% to RhCMV 180.92 strain (691). To further characterize this novel virus, I compared and contrasted the structural and functional genes of CyCMV to human and rhesus macaque CMV with respect to pathogenesis, immune evasion, and species-specificity.
Furthermore, the genomic sequencing and characterization of CyCMV facilitated the identification of target sites required for cloning the virus as a bacterial artificial chromosome.

In order to insert the vaccine antigens into the CyCMV vector, the linear virus must be cloned into a plasmid. However, like all herpesviruses, the extremely large size of the viral genome prevents traditional cloning and requires the use of advanced cloning methodologies, such as bacterial artificial chromosomes (BACs), which are capable of sustaining large DNA inserts. The Cre/loxP recombination system was used to facilitate the cloning of CyCMV as a BAC, which required the insertion of a loxP recognition sequence into viral genome. To this end, two recombinant viral intermediates were generated, CyCMV-eGFP and CyCMV-loxP. To create CyCMV-eGFP, an eGFP marker flanked by two loxP sites was inserted into the viral genome by homologous recombination. Subsequently, to generate a single loxP site for BAC recombination, the eGFP marker was excised and the loxP sites were recombined by Cre/loxP recombination to form a single recognition sequence in CyCMV-loxP. The BAC plasmid encoding an eGFP selection marker was cloned into the viral genome by Cre/loxP recombination to successfully generate a CyCMV-eGFP-BAC recombinant virus. The newly circularized virus was transferred into a bacterial system, which will allow for downstream viral manipulation and the insertion of vaccine antigens for the development of CyCMV as an HIV/SIV vaccine.

Moreover to our primary interest in CyCMV as an HIV vaccine vector, this newly identified endogenous virus will have a variety of applications in infectious disease and transplant research, in which this species of macaques is widely utilized. We hope this newly sequenced and characterized CyCMV genome will provide the necessary groundwork for future studies evaluating the utility of cynomolgus macaques as a NHP model in which to study HCMV vaccine development, CMV biology and associated pathologies. Congenital CMV remains the most common viral cause of birth defects in newborns and yet there is still no vaccine (113). The burden of CMV disease is not only apparent in children but also in adults, specifically those receiving solid organ or bone marrow transplants, and those suffering from an immune compromising disease, such as HIV/AIDS. We have reason to be hopeful regarding the ability to make a CMV vaccine given the success of another herpesvirus, Varicella Zoster Virus (VZV), in which the licenced vaccine has been highly effective in reducing the mortality associated with varicella infections in the United States (503). Furthermore, with the potential for dual CMV/VZV-vectored HIV vaccination, we propose to perform a side-by-side/concurrent
evaluation of VZV and CMV to compare and contrast the immune responses exhibited by these two highly similar, yet very distinct herpesviruses (outlined in Table 1-1: Comparison of VZV and CMV as HIV/SIV Vaccine Vectors). With the unique features of these two herpesviruses, it is possible that they may work in combination to improve the quantity and quality of the anti-HIV/SIV immune response to achieve enhanced protective efficacy beyond what has currently been observed with CMV vaccination alone in the rhesus macaque model.

5.1.2 Rhesus Macaque Cytomegalovirus-Based Viral Vector

During the time the work described in this thesis was being conducted, Louis Picker’s group at the Oregon Health and Science University published very promising results on their efforts developing rhesus macaque CMV (RhCMV) as an HIV viral vector. We have been fortunate to be able to learn from the successes and failures of the RhCMV vaccine studies throughout the development of CyCMV as an HIV/SIV viral vector. Since 2009, two large-scale rhesus macaque vaccine trials have been published and will be described in detail. In the first efficacy trial, rhesus macaques were immunized with three different RhCMV 68.1 BAC-derived vectors expressing SIVmac239 proteins Gag, a chimeric antigen Retanef (Rev, Tat, Nef), and Env, respectively, over an extended vaccination schedule (338). The three vectors were consecutively given in 19 week intervals followed after by a tri-vector combination boost at week 59 (Figure 5-1). The vaccine-induced an accumulation of SIV-specific CD4+ and CD8+ T cells of effector memory phenotype that expressed antiviral cytokines (TNFα, IFNγ), chemokines (MIP-1β), and had increased surface expression of the cytolytic-associated CD107a degranulation marker. These vaccine-generated effector memory-differentiated T cells were present and maintained in high frequency at mucosal sites (Bronchial Alveolar Lavage) for up to 8.4 years after the initial vaccination (334). In addition, the vaccine was well tolerated and productive infection was maintained in the vaccinated animals, as demonstrated by RhCMV-SIV shedding in the urine and saliva for almost two years following inoculation. Greater than 69 weeks after the vaccination boost, the animals were challenged intrarectally with low-dose mucosal SIVmac239 at repeated weekly doses of 300 focus-forming units (FFU) for 8 weeks followed thereafter by 1000 FFU for 5 weeks, at which point all animals were productively infected with SIV (Figure 5-1). Following SIV infection, 4 of the 12 vaccinated animals did not display progressive infection as demonstrated by low or undetectable plasma viral load, which
was confirmed by CD8\(^+\) T cell depletion in which SIV DNA and RNA remained undetectable by sensitive PCR assays (338).

Figure 5-1 RhCMV-SIV Vaccine Efficacy Trial. [Reprinted with permission from Macmillan Publishers Ltd: Nature Medicine, (338), copyright 2009]

A second, larger vaccine trial was conducted to build on these results by modifying the vaccine constructs and the vaccination regimen to include a prime-boost vaccination strategy (334). The first group was primed and boosted with a cocktail of four RhCMV-SIV vectors, which included the three previously examined RhCMV-Gag, RhCMV-Retanef, RhCMV-Env, and a novel RhCMV vector encoding SIV Pol (RhCMV-Pol) (Figure 5-2). In addition to evaluating the protective efficacy of RhCMV-SIV vectors, this trial was designed to also compare the immunogenicity, T cell phenotype bias, and efficacy of a CMV-based vector versus replication-defective Ad5-SIV vector. Two additional groups were included with a heterologous prime/boost vaccination in which the animals received a single prime with RhCMV-SIV cocktail, or three monthly primes with a DNA vaccine expressing pan-SIV proteome, and both groups were boosted with an Ad5 vector expressing pan-SIV proteome (Figure 5-2). For all three groups, the boosts were given 14 weeks after the first prime. The primary difference between the vaccination strategies was the vaccine-induced SIV-specific CD8\(^+\) T cell memory phenotype, with the RhCMV-SIV/RhCMV-SIV prime/boost eliciting an effector memory T cell response compared to the DNA/Ad5 strategy, which induced a central memory-biased CD8\(^+\) T cell response (334). This highlights the limitations of the HIV vaccine vectors (Adenovirus and Poxvirus) currently being tested in clinical trials, which generate delayed and transient HIV-specific central memory-derived T cell responses, as described in Section 1.2.3.4 (Ongoing HIV Vaccine Clinical Trials). Furthermore, in the RhCMV-SIV/RhCMV-SIV vaccinated animals, SIV Gag-specific memory T cells responses were maintained in high frequency with a broad tissue distribution, including gastrointestinal mucosal tissues, for over 2 years post-vaccination (334). These results support our overarching hypothesis for using CyCMV as an HIV/SIV viral vector with the potential to induce a robust and broad effector memory-biased T cell response at
the portal of HIV/SIV entry (reviewed in Sections 1.3.1 Rationale for Herpesvirus-Based Viral Vectors, and 1.3.4 Overarching Hypothesis).

Figure 5-2 RhCMV-SIV Prime-Boost Vaccine Trial Schedule. [Reprinted with permission from Macmillan Publishers Ltd: Nature, (334), copyright 2011]

Forty-five weeks after the boost, the animals were challenged with a multi-low dose SIVmac239 intrarectal challenge until all animals were infected (Figure 5-2) (334). In their analysis of the results, the outcomes in the animals from the groups that were primed with RhCMV-SIV, were similar regardless of the type of boosts that were given (RhCMV-SIV or Ad5), and therefore the two groups were pooled. Thirteen of the twenty-four (~54%) RhCMV-SIV primed animals were able to control infection (termed “controllers”), marked by undetectable plasma viral load with sporadic blips that decreased in frequency overtime, suggesting that the virus was eventually cleared (Figure 5-3). As a marker to differentiate vaccine- versus challenge virus-induced SIV-specific T cell responses, Vif protein was not present in the vaccine constructs but is naturally expressed in the SIVmac239 challenge virus. The controllers had persistent SIV-specific T cell responses to the antigens incorporated in the vaccine (Gag, Pol), while the Vif-specific T cell responses waned overtime, further demonstrating viral clearance. None of the DNA/Ad5 vaccinated animals or unvaccinated controls were able to control infection, though the viral load was reduced in the DNA/Ad5 animals compared to the unvaccinated controls (Figure 5-3). In an effort to confirm viral clearance in the RhCMV-SIV vaccinated controllers, CD8+ and CD4+ T cell depletions were performed. The plasma viral load did not increase in any of the controllers and Vif-specific T cell responses remained absent, suggesting viral clearance was achieved in this subset of vaccinated animals. After over 1 year of follow-up, 12 of the 13 controllers were able to maintain
control of SIV infection. At necropsy, an extensive examination of SIV in the gastrointestinal tract, GALT, peripheral lymph nodes, and secondary lymphoid tissues was performed and revealed the presence of detectable cell-associated SIV RNA and DNA in 28% and 20% of the samples, respectively, however no detectable replicating SIV was recovered by co-culture. In an analysis of the correlates of protection, a significant correlation was observed between the burst size of the peripheral blood SIV-specific CD8$^+$ memory T cell response following the vaccine boost and the animals that controlled SIV infection (334). This correlation suggests that the magnitude of the systemic immune response to vaccination may determine the number of vaccine-induced SIV-specific effector memory T cells present at the effector sites that are able to control SIV replication and inhibit viral dissemination.

![Figure 5-3 RhCMV-SIV Prime-Boost Vaccine Trial Results](image)

**Figure 5-3 RhCMV-SIV Prime-Boost Vaccine Trial Results.** [Reprinted with permission from Macmillan Publishers Ltd: Nature, (334), copyright 2011]

Given that T cell-based vaccine strategies are largely aimed at reducing HIV/SIV viral load and modifying disease course, it is not surprising that sterilizing immunity was not achieved and neutralizing antibodies were not detected in these studies. In both trials, there was no significant difference between groups in the number of SIV challenges required to achieve infection however, there was a striking difference in the number of vaccinated animals that controlled SIV infection, although it is not known how to predict which animals will respond and which animals will not (334, 338). Furthermore, the RhCMV vectors expressing SIV Env did not elicit neutralizing SIV-specific antibodies, though low-titre nonneutralizing anti-SIV antibodies were observed in the first efficacy trial (334, 338). During the vaccination trial, it became clear that the RhCMV-SIV vaccine-induced SIV-specific T cell responses were targeting different epitopes than the responses induced by natural infection or conventional vaccination (334). This led Louis Picker’s group to evaluate these responses in greater detail. It was observed that the
majority of the RhCMV-SIV-vected effector memory CD8^+ T cell responses were not directed towards the canonical MHC class I-restricted immunodominant and well-characterized SIV epitopes within Gag (CM9) and Tat (SL8), but were in fact predominantly targeting completely different, and nonconventional SIV epitopes that were MHC class II-restricted (336). The RhCMV-SIV-induced Gag-specific CD8^+ T cells responses had a 3-fold increase in the breadth of recognition compared to conventional vaccine responses. This study demonstrated that this phenomenon was unique to the fibroblast-adapted RhCMV 68.1 strain, which lacks the HCMV homologues of UL128-131 involved in cellular tropism. They propose that the MHC class I downregulation gene, US11 mediates these unconventional MHC class II-restricted CD8^+ T cell responses, however US11 expression is typically suppressed by UL128-131 and therefore the absence of these genes results in atypical epitope recognition (336).

Interestingly, CyCMV (Ottawa and Mauritius strains) are also fibroblast-adapted viruses, however they are relatively low-passaged viruses that have retained the cellular tropism encoded HCMV US128-131 homologues (as described in Section 3.5.5 Tropism Genes, and Table 3-1). Thus, in contrast to RhCMV, the CyCMV encoded US128-131 homologues would likely suppress US11-mediated MHC class II-restricted T cell responses, though the phenotype of the T cell responses would have to be characterized in SIV-infected cynomolgus macaques. It appears that the ability of SIV-specific CD8^+ T cells to recognize MHC class I- and II-restricted epitopes would be advantageous, however it unclear if the unique epitope targeting of RhCMV-vected responses contributes to the control of SIV infection observed in their studies. Another unique difference between CyCMV and RhCMV is that CyCMV does not encode a viral cyclooxygenase-2 (COX-2) gene that appears to be unique to RhCMV (337). Cellular COX-2 expression is induced upon HCMV infection and has been shown to play an important role in HCMV replication (914). Unlike HCMV infection, RhCMV infection does not induce cellular COX-2 expression in the presence of the viral COX-2 isoform encoded in the RhCMV genome (rh10) (708). Future studies will be required to determine if CyCMV infection induces cellular COX-2 expression in the same way as HCMV.

5.1.3 Limitations and Safety Concerns to CMV-Based Vectors

In light of the many failed attempts at generating an effective HIV vaccine, the HIV/SIV vaccine field has returned to examining live attenuated viruses in NHPs, as the risk-benefit ratio
of this pandemic is driving a willingness to examine any and all avenues to establish protective immunity. We acknowledge that CMV is a pathogen, however it is typically asymptomatic in the general population, and is only harmful in individuals with an immature immune system or who are immunocompromised. That said, efforts would be made to attenuate CMV-based HIV viral vectors in an attempt to reduce or prevent any vaccine-associated morbidity. In this regard, we will be able to build on previous efforts to generate a preventative HCMV vaccine. Specifically, the live attenuated Towne 125 strain was not successful at preventing HCMV infection as the degree of attenuation was too robust to mount a sterilizing immune response, however the vaccine was effective at preventing or reducing disease in immunosuppressed individuals (8, 300, 643, 647-650). Another important, yet controversial topic related the safety of CMV-based vectors is the relationship between CMV seropositivity and its influence on mortality risk in elderly individuals (625, 880). A number of researchers are striving to either prove or disprove correlations between chronic CMV infection and increased morbidity or premature mortality (7, 217, 589, 862). Much of the debate is centered on the relationship between CMV seropositivity in the elderly and the associated Immune Risk Profile (IRP), which includes an altered CD4⁺/CD8⁺ T cell ratio, increases in the number of terminally-differentiated CD8⁺ T cells, and reduced T-cell proliferation (135, 148, 606, 876). Furthermore, the relationship between the level of CMV IgG antibody titres and CMV-associated pro-inflammatory cytokines has also been evaluated in this context (693, 795, 862). Given the high seroprevalence of CMV, particularly in the regions most devastated by HIV/AIDS where the prevalence is close to 100%, the benefits certainly outweigh potential risks, however further investigation into the proposed causal link between CMV infection and mortality in the elderly population is of great importance.

Efforts have been made to improve the safety profile of CMV by enhancing immune control over CMV infection and attenuating the virus to reduce pathogenicity. CMV encodes a number of immune evasion genes that allow the virus to evade the immune response and establish lifelong infection in the host. Thus, enhanced host immune control over the virus could serve to improve the safety characteristics of CMV as a viral vector. Initial attempts to remove the MHC class I immunoevasins (homologues of HCMV US2-US11) from RhCMV resulted in primary infection in seronegative rhesus macaques however, in RhCMV-seropositive animals, RhCMVΔUS2-11 was not able to overcome the RhCMV-specific CD8⁺ T cell response, thus preventing superinfection (335). Given the high prevalence of CMV, it is imperative that the
vaccine has the capacity to re-infect CMV-positive individuals. In a similar regard, CMV is horizontally transmitted through bodily secretions and as such the viral vector could be transmitted to unvaccinated individuals. In order to prevent transmission of the vaccine vector, particularly to immunocompromised individuals, CMV-based vectors should be modified and/or attenuated. Louis Picker’s group is currently evaluating the removal of a tegument protein, pp71 (Rh110), from RhCMV and early studies have demonstrated that the attenuated spread-deficient virus was not shed in the urine of RhCMVΔpp71 vaccinated rhesus macaques (unpublished; Keystone Conference, Whistler, Canada, 2011). Furthermore, murine CMV (MCMV) attenuation studies have demonstrated that both spread-deficient MCMV (MCMVΔgL), and single cycle MCMV strains establish latency and maintain the ability to generate memory inflated CMV-specific effector memory CD8⁺ T cells (769). Another MCMV attenuation study examined the effect of enhancing immune control of the virus by generating a recombinant MCMV-expressing an NK cell activating receptor (NKG2D) (767). The results of this study demonstrated that inoculation with the attenuated virus was immunologically controlled in both immunocompetent and immunocompromised mice, and notably, it was associated with reduced CMV-associated disease. Moreover, the attenuated virus retained the ability to mount a potent CMV-specific immune response, which was protective from subsequent MCMV infection with a lethal challenge strain (767). The robust CMV-specific effector memory response appears to be independent of viral replication, thus an attenuated CyCMV may improve the pathogenenicity of the vaccine while retaining the immunogenicity and allowing for persistent immune presentation of the vaccine antigens. These strategies, as well as future evaluations, will need to be further evaluated in NHP models in an effort to provide the necessary groundwork to ensure that CMV will present a safe and effective viral vector for an HIV vaccine to be evaluated in humans. Ultimately, a co-HIV/HCMV vaccine that could prevent both HIV and HCMV acquisition would be the most favourable outcome, however a significant amount of research and development, as well as safety evaluations would be required before this could theoretically be achieved.

5.2 Future Directions

5.2.1 CyCMV-eGFP-BAC Confirmation and Viral Reconstitution

Having cloned CyCMV as a BAC, a number of confirmation and evaluation steps will need to be performed prior to proceeding to the insertion of the vaccine antigens. These future
directions have been largely discussed in Section 4.6 (Chapter 4 Discussion), thus I will briefly expand on particular confirmation experiments. To confirm the location of the BAC insert, a Southern blot experiment has been designed with hybridization probes specific for cyUS1, cyUS2, loxP, and eGFP. These probes will serve to confirm the location of the inserted DNA in specific fragments generated by restriction digestion, with NheI and XbaI, of the recombinant viruses (CyCMV-eGFP, CyCMV-loxP, CyCMV-eGFP-BAC) and the parental strain (CyCMV), as a control. Upon confirming the inserts, next-generation sequencing and genetic characterization of the CyCMV-eGFP-BAC plasmid will provide insight into which, if any, CyCMV genes were disrupted during the cloning process. To ensure the recombinant virus has retained its infectiousness following propagation in bacteria, the CyCMV-eGFP-BAC plasmid will be transfected into a fibroblast cell line to assess viral reconstitution and in vitro propagation. Upon successful confirmation of the cloned virus, the vaccine antigens will be incorporated to generate the CyCMV vaccine constructs.

5.2.2 Development of CyCMV-Based HIV/SIV Vaccine Candidates

5.2.2.1 Vaccine Antigens

In terms of the immunogens to be incorporated into the vector, we chose to pursue two distinct approaches. The first was a novel unproven approach, focused on the use of endogenous retroviral genes as targets for the immune response. The second approach is a more conventional strategy that utilizes a variety of SIVmac239 genes, and it is hoped that a CMV-based delivery system will enhance the effectiveness of these antigens in a vaccine setting.

5.2.2.1.1 Novel Antigens

The work described in this Section has been previously published in PLoS ONE, 2012, 7 (6) (Appendix 1) (505). One unique vaccine antigen approach that is under investigation is to examine HIV-infected cells to determine if the virus causes changes to cellular proteins that may act as surrogate markers for HIV infection and thus represent novel vaccine targets. Recently, human endogenous retroviruses (HERVs) have emerged as potential alternative cellular targets for this type of HIV vaccine strategy (288, 738). HERVs are remnants of ancient retroviral infections, fixed in our genome and transmitted vertically (847). Integration of HERVs has occurred over millions of years and over time they have acquired point mutations rendering them unable to produce infectious virions (483, 585). The relationship between HERV and HIV
emerged from initial reports demonstrating that antibodies against HERV were found in 70% of HIV-positive patients compared with only 3% of HIV-naïve individuals (214, 483). Furthermore, it has been shown that HERV RNA titers are elevated in the plasma of HIV-infected individuals in comparison with seronegative individuals (173-175, 288). Most intriguing for vaccine design is a recent study demonstrating elevated and functional T-cell responses in HIV-positive patients against cellular protein targets derived from HERV (288). Given that HERVs are encoded in the germ-line, it is hypothesized that they are not subjected to the same degree of cytotoxic T lymphocyte (CTL) immune pressure and subsequent immune escape as HIV antigens.

To investigate this further, our group conducted a longitudinal study using a non-human primate model, which allowed for endogenous retrovirus (ERV) expression to be assessed pre- and post-SIV infection from dependent samples, providing a precise measurement of changes in ERV expression as a result of viral infection. I evaluated cynomolgus macaque endogenous retrovirus (CyERV) mRNA expression in PBMCs following inoculation with a live attenuated SIV vaccine, and subsequent SIV infection (878). Our findings were not consistent with what has been observed for HERV expression in the plasma of HIV-infected humans. CyERV Env and Gag mRNA expression was decreased following acute SIV-infection, whereas during chronic SIV infection, CyERV transcript levels were indistinguishable from baseline (Figure 5-4).

![Figure 5-4 CyERV Gene Expression in Peripheral Blood Mononuclear Cells](505). CyERV Envelope and Gag RNA was isolated from PBMCs and quantified using RT-qPCR. Log transformed CyERV Envelope (A) and Gag (B) gene expression levels relative to GAPDH were
compared across all timepoints (baseline, post-vaccination, acute and chronic SIV infection). Vaccine groups include controls, Δ5-CMV group, and Δ6-CCI group. Paired-samples t-tests (2-tailed) were performed to determine statistical significance, bars represent the mean.

Furthermore, *in vitro* analysis revealed that SIV infection of purified CD4+ T cells did not alter CyERV gene expression. This study represented the first evaluation of ERV expression in cynomolgus macaques following SIV infection and it is evident that future examination of CyERV expression in plasma and tissue samples would assist in further explaining whether these findings are unique to macaques or if this is a cellular phenomenon that has yet to be fully investigated in humans. Thus, further investigation at the cellular level will be required to elucidate the impact of HIV/SIV infection on endogenous retrovirus expression. A study was recently published examining the safety and immunogenicity of vaccinating rhesus macaques with simian endogenous retrovirus (SERV) Gag and Env genes, which were delivered with a DNA prime and recombinant Adenovirus subtype 5 (Ad5) boost regimen (717). The results of this study demonstrated that the SERV genes were able to elicit anti-SERV Gag and Env T cell responses, as well as anti-SERV Env antibody responses in these macaques. This group went on to challenge these vaccinated macaques with SIV in an effort to evaluate the efficacy of the vaccine, however the results of the challenge phase have not yet been published. It will be interesting to see how the rhesus macaque SIV challenge model compares to our results, as our non-human primate model system did not recapitulate what has been observed to date in the plasma of HIV-infected humans. Furthermore, it is important to acknowledge that evaluating the efficacy and feasibility of this concept in non-human primate models may not translate to humans, and this may still be a viable HIV vaccine strategy. In this regard, further investigation of human endogenous retrovirus expression in a variety of tissues derived from HIV-negative and HIV-positive humans will provide more information on the utility of this concept for an HIV vaccine. However, in an effort to proceed with our focus of generating CyCMV vaccine constructs, conventional SIV antigens will be evaluated at this time.

### 5.2.2.1.2 Conventional SIV Antigens

Although the use of SIV genes for SIV vaccines is a conventional approach, it is hoped that the features of a CMV-based vaccine vector described in Section 1.3.1 (Rationale for Herpesvirus-Based Viral Vectors) will augment the immune response allowing the SIV antigens to elicit broad cellular and humoral immune responses with the potential benefit of persistent
antigen exposure. In choosing the particular antigens to be utilized in an HIV/SIV vaccine, countless iterations have been tried, from single-gene to multi-gene constructs, but the precise combination required to elicit the greatest degree of protective or therapeutic effect has yet to be empirically determined. In an attempt to induce broad and long-lasting cellular and humoral immunity, we have employed both structural (Gag, Pol, Env) (Figure 5-5A), and non-structural components comprised of transactivator (Tat, Rev), and regulatory (Nef) genes (Figure 5-5B). The CD8$^+$ T cell targeted proteins Gag, Pol, Nef, Tat, and Rev are presented early after virus infection, thus enabling CD8$^+$ T cell-specific responses to be generated during the acute phase of infection (718). While the Env protein, generally targeted by CD4$^+$ T cells to produce antibodies, is expressed late in the viral replication cycle. It is hoped that this combination of SIV genes will allow for a multi-faceted immune response.

**Figure 5-5 Schematic of SIVmac239 GFE and NTR Expression Constructs.** Codon-optimized versions of SIVmac239 Gag, Pol, Env and a novel fusion protein NeTaRev, comprised of Nef, Tat, and Rev were generated by de novo PCR-assembly. The size of each individual protein-coding region is indicated along with the complete length of each construct. Both GFE and NTR expression constructs were driven by the pCMV immediate early promoter and flanked
downstream by a bovine growth hormone poly-adenylation signal (Bgh pA). A) Within the GFE expression cassette, the ribosomal frameshift present within the Gag-Pol polyprotein was removed (ΔFS) to generate a Gag-Pol fusion. Translation of Env is facilitated by an internal ribosomal entry site (IRES) element. B) NeTaRev is a fusion protein derived from codon-optimized derivatives of SIVmac239 Nef, Tat, and Rev. The C- and N358 terminal halves of Tat have been juxtaposed. (Adapted from Willer, D. O., Marsh, A. K., et al. unpublished).

Codon-optimized SIV antigen cassettes have been previously prepared by the MacDonald laboratory and are readily available for evaluation in a CyCMV vectored vaccine. All SIV genes are derived from SIVmac239, the highly pathogenic clone of SIV from sooty mangabeys and passaged in rhesus macaques (27) (discussed in Section 1.1.1 HIV/SIV Structure and Genome). Two variants of the SIVmac239 antigen cassettes have been constructed with the first being SIVGFE, which will allow for expression of a Gag-Pol fusion protein, and the Env protein (Figure 5-5A). The other SIV antigen cassette is termed SIVNTR, which is a novel fusion protein comprised of Nef, Tat, and Rev coding sequences (Figure 5-5B). Our group has inserted these SIV antigenic cassettes (GFE and NTR) into human VZV for evaluation in a VZV-based SIV vaccine trial (Willer, D. O., Marsh, A. K., et al. unpublished). With the insertion of GFE (8,588 bp) and NTR (2,495 bp) into the VZV genome, it was important to evaluate viral growth kinetics to ensure the antigens have not hindered viral fitness. In this regard, I confirmed that the recombinant VZV viruses carrying the GFE and NTR expression cassettes (VZV-GFE, VZV-NTR) were indistinguishable from the wild-type virus (VZV-pOKA) across all time points examined (Figure 5-6) (Willer, D. O., Marsh, A. K., et al, unpublished). Thus, the introduction of GFE and NTR into the VZV genome had no effect on the in vitro growth properties of the virus. Furthermore, with respect to CMV-based vectors, it was also confirmed by Louis Picker’s group that the insertion of comparable SIV antigens into RhCMV did not impede viral growth kinetics (338).
Figure 5-6 Viral Growth Kinetics of VZV-SIV Vaccine Constructs. VZV-pOKA, VZV-GFE and VZV-NTR were grown on MeWo cells and harvested daily for 5 days post-infection. Viral titres were determined using a standard VZV plaque assay. Data points on the graph represent the mean viral titre (Log$_{10}$ Plaque Forming Units/ml) of duplicate assays from two independent experiments with the error bars representing the Standard Error of the Mean (SEM). (Willer, D. O., Marsh, A. K., et al. unpublished).

To prepare our VZV-focused SIV vaccine constructs for insertion into CyCMV, a number of adaptations need to be performed. Efforts are currently underway in the MacDonald laboratory to modify the constructs to encode elongation factor-1 alpha (EF-1 alpha) promoters in place of the CMV promoters, and to generate GFE and NTR containing plasmids encoding the necessary flanking CyCMV genes, which will be the target location for the insertion of the vaccine antigens into CyCMV-BAC via allelic exchange bacterial recombination to generate CyCMV-BAC-SIV. After establishing appropriate growth, protein expression of the CyCMV-BAC-SIV vaccine will be evaluated for transgene stability and gene expression by Western blot. Furthermore, the immunogenicity of the vaccine antigens will be examined by flow cytometry and ELISA, in order to evaluate anti-SIV cellular and humoral immune responses, respectively. Much of the work described herein is part of a large team grant in which other team members are working to complement the protective efficacy of a T cell-based CyCMV-vectored HIV/SIV vaccine by examining strategies that will be tested in conjunction with the CyCMV vector to improve the anti-HIV/SIV humoral response at the mucosa. Such approaches involve the use of
novel B cell immune adjuvants, which include Nod-Like Receptor (NLR) agonists and/or members of the Tumor Necrosis Factor SuperFamily (TNFSF). In addition, efforts are underway in the laboratories of our collaborators, Dr. Mario Ostrowski and Dr. James Rini, to develop a novel vaccine antigen by generating a trimeric fusion protein comprised of TNFSF proteins involved in B cell immunity, such as BAFF or APRIL, and a trimeric HIV-1 Env gp140 protein (termed TNFSF-HIV Env). It is hoped that these approaches aimed at augmenting vaccine-induced B cell immunity will be successful at generating a robust and protective antibody response, either mediated by ADCC, or ideally a broadly neutralizing antibody response that will work synergistically with the persistent CyCMV vector to induce sterilizing immunity and prevent HIV/SIV infection.

5.2.2.2 Generation of Self-Excisable BAC

Following the successful insertion and confirmation of the vaccine antigens, the BAC backbone will be removed from the CyCMV-eGFP-BAC-SIV vaccine constructs prior to in vivo testing. To facilitate the removal of the cloning plasmid, a self-excisable BAC will be generated by replacing the eGFP gene located within the BAC plasmid with Cre, which will be executed by allelic exchange to generate CyCMV-Cre-BAC-SIV (128, 903). The resulting plasmid will then be transfected into mammalian fibroblast cells allowing for in vitro transcription of the Cre recombinase enzyme, which mediates Cre-loxP recombination in the two loxP sites encoded in CyCMV-Cre-BAC-SIV. This recombination will facilitate the removal of the BAC vector from the viral genome, leaving only a single 34bp loxP sequence in the reconstituted progeny virus (rCyCMV-SIV). The fidelity of the Cre-loxP excision of the BAC sequence will require verification by Southern blot and DNA sequencing across the site of excision. At this stage, the infectious CyCMV-SIV vaccine candidates will be ready for evaluation in a cynomolgus macaque model.

5.2.3 Proposed CyCMV-SIV Vaccine Trial

The first step for in vivo evaluation of the CyCMV-SIV vaccines will be to assess the immunogenicity of the vaccines in cynomolgus macaques. Briefly, adult CyCMV-seropositive cynomolgus macaques will be assigned in a pair-wise manner to account for MHC haplotype into two groups, receiving inoculations of either CyCMV-SIV or control (CyCMV vector or media). The vaccine group will receive a vaccine cocktail comprised of CyCMV-SIV_{GFE} and
CyCMV-SIV\textsubscript{NTR}, as well as any B cell immune adjuvants that have been incorporated (as discussed in Section 5.2.2.1.2 Conventional SIV Antigens), while the controls will receive the CyCMV vector backbone alone or media. Having reviewed multiple HIV vaccination strategies as well as the most successful NHP and human HIV/SIV vaccine trials conducted to date (Section 1.2 HIV/SIV Vaccine Development), I would propose a prime-boost vaccination strategy in an effort to achieve optimal cellular and humoral immunogenicity of the vaccine. The animals would receive two subcutaneous primes at 12-week intervals with CyCMV-SIV and would be boosted thereafter with the proposed fusion protein (TNFSF-SIV Env). In order to determine if the vaccine was well tolerated and safe, routine health exams, hematology, and clinical monitoring will be performed. To evaluate immunogenicity, a variety of tissue samples will be collected throughout the study including PBMCs, urine, saliva, broncho-alveolar lavage (BAL), lymph nodes biopsies, genital swabs, rectal swabs, and will be evaluated by assessing the cellular and humoral immune responses to the SIV antigens. Anti-SIV cellular responses will be stimulated with in-house SIVmac239 peptide libraries and assessed by flow cytometry. The MacDonald laboratory has previously designed and optimized flow cytometry panels to enumerate, and evaluate the polyfunctionality of various cellular populations, including memory subsets, CD4\textsuperscript{+} T cell subsets (Th17, Th22, Tregs), CTLs, as well as assess activation, proliferation, and the induction of multiple cytokines. ELISA, Western blot, and SIV neutralization assays will be used to evaluate antibody responses generated by the vaccine. Based on the results of the immunogenicity trial, modifications to the vaccination regimen, sampling schedule, or outcome measurements may be required prior to continuing with the challenge trial.

If the vaccine is deemed safe and immunogenic, a multi-low dose mucosal SIV challenge trial will be conducted to evaluate the efficacy of the vaccine in controlling SIV infection. The vaccine and control groups will be the same as described above, with the media control group serving to delineate the potential impact of CyCMV inoculation on SIV acquisition or pathogenicity. Animals will receive weekly intra-rectal heterologous challenges with an escalating low dose regimen of pathogenic SIVmac251. The animals will be continuously challenged until seroconversion is detected by RT-qPCR. Weekly tissue sampling will be performed, as described above in the immunogenicity study, and SIV viral loads will be quantitated weekly by RT-qPCR to follow SIV disease progression. Similar to the experiments describe in the immunogenicity trial, post-SIV infection monitoring will be conducted and will
include immunophenotyping of the peripheral and mucosal samples to evaluate the immunogenicity and efficacy of the vaccine, as well as monitoring disease progression and the clinical status of the macaques. For flow cytometric analysis, samples will be stimulated with either SIV Gag peptide pools, or SIV “other” peptide pools that comprise SIV genes not present in the vaccine vector (i.e. Vif, Vpx, Vpr), which will serve to differentiate between vaccine- and challenge-induced cellular responses. In addition to the experiments briefly described in the immunogenicity study, a collaborative effort between experts across North America will be required to appropriately analyze mucosal and serum antibody responses including anti-SIV IgG, IgA, ADCC, ADCVI, and broadly neutralizing antibodies. Furthermore, a systems biology approach could be used to examine the RNA transcriptome in specific cell types in an effort to gain further insight into the profile of vaccine-induced changes in gene expression. Immunohistochemistry of various necropsy samples will allow for tissue-specific visualization of changes in the architecture of secondary lymphoid tissues, as well as characterization of resident immune cells and cellular cross-talk. Depending on the outcome of the study, for example, if animals are controlling SIV infection then CD8\(^+\) T cell depletion studies, viral co-culture experiments, along with highly sensitive nested PCR or digital droplet PCR in multiple tissues obtained at necropsy, may be performed to demonstrate viral clearance.

The primary endpoint of the CyCMV-SIV vaccine/SIV challenge trial will be vaccine-induced sterilizing immunity and the prevention of SIV infection. If the animals become infected, the secondary outcome will be to evaluate the impact of vaccination on SIV disease course and outcome through the examination of modifications in peak and set-point viral loads, along with the rate of CD4\(^+\) T cell depletion, and progression to simian AIDS. It is hoped that the results will provide informative outcomes that will allow for further optimization of a number of parameters, including but not limited to, the vaccination regimen, immune adjuvants, and tissue sampling in order to be re-evaluated in subsequent non-human primate studies. Ultimately, the goal of these studies is to provide translational results that will serve to validate the tolerability, safety, and efficacy of this vaccination strategy for evaluation in human clinical trials.

5.3 Conclusion

The use of CMV as a viral vector for an HIV vaccine shows much promise based on its ability to elicit a robust and persistent effector memory CTL response at mucosal site. We
propose that a CMV-based vector will prime the immune response to eliminate HIV immediately upon entry and prior to viral dissemination, thus preventing systemic infection and reducing viral burden. In an effort to evaluate the safety and effectiveness of CMV as a viral vector prior to human clinical trials, we established a non-human primate model using cynomolgus macaques. Given the strict species-specificity of CMVs, it was obligatory to use cynomolgus macaque-specific CMV. We were the first to isolate, characterized, and sequence the viral genomes of CyCMV from animals of Filipino and Mauritius origin. A thorough comparison between CyCMV, HCMV, and RhCMV was performed with particularly focus on the functional genes involved in pathogenesis, immune evasion, and species-specificity. The genomic sequence of the virus was used to direct the cloning of CyCMV as a BAC. CyCMV-BAC was generated in order to maximize the ability to manipulate the virus and facilitate the insertion of vaccine antigens. The work described herein provides the necessary groundwork to evaluate a CyCMV-based SIV vaccine in a cynomolgus macaque model, with the ultimate goal of providing evidence of protective efficacy and safety that will inform the development of an HCMV-based HIV vaccine to evaluate in human clinical trials.
Appendices

Evaluation of Cynomolgus Macaque (Macaca fascicularis) Endogenous Retrovirus Expression Following Simian Immunodeficiency Virus Infection

Angie K. Marsh1,2,3, David O. Willer2,3, Olena Skokovets2, Oluwadamilola H. Iwajomo2,3, Jacqueline K. Chan2,3, Kelly S. Macdonald1,2,3*

1 Department of Immunology, University of Toronto, Toronto, Ontario, Canada, 2 Department of Medicine, University of Toronto, Toronto, Ontario, Canada, 3 Department of Microbiology, Mount Sinai Hospital, Toronto, Ontario, Canada

Abstract

Human endogenous retrovirus type K (HERV-K) transcripts are upregulated in the plasma of HIV-infected individuals and have been considered as targets for an HIV vaccine. We evaluated cynomolgus macaque endogenous retrovirus (CyERV) mRNA expression by RT-qPCR in PBMCs isolated from a cohort of animals previously utilized in a live attenuated SIV vaccine trial. CyERV env transcript levels decreased following vaccination (control and vaccine groups) and CyERV env and gag mRNA expression was decreased following acute SIV-infection, whereas during chronic SIV infection, CyERV transcript levels were indistinguishable from baseline. Reduced susceptibility to initial SIV infection, as measured by the number of SIV challenges required for infection, was associated with increased CyERV transcript levels in PBMCs. In vitro analysis revealed that SIV infection of purified CD4+ T-cells did not alter CyERV gene expression. This study represents the first evaluation of ERV expression in cynomolgus macaques following SIV infection, in an effort to assess the utility of cynomolgus macaques as an animal model to evaluate ERVs as a target for an HIV/SIV vaccine. This non-human primate model system does not recapitulate what has been observed to date in the plasma of HIV-infected humans suggesting that further investigation at the cellular level is required to elucidate the impact of HIV/SIV infection on endogenous retrovirus expression.

Introduction

Despite major advances in antiretroviral therapy, the HIV/AIDS epidemic continues to spread. Researchers are striving to prevent HIV acquisition through vaccination, however the very nature of HIV poses significant challenges to the development of a successful HIV vaccine. HIV rapidly mutates and recombines under immune selection pressure leading to extreme antigenic diversity and hypervariability (reviewed in [1,2]). Conventional vaccine approaches have been largely ineffective against this degree of viral diversity. One unique approach under investigation is to examine HIV-infected cells to determine if the virus causes changes to cellular proteins that may act as surrogate markers for HIV infection and thus represent novel vaccine targets. Recently, human endogenous retroviruses (HERVs) have emerged as potential alternative cellular targets for this type of HIV vaccine strategy [3,4]. HERVs are remnants of ancient retroviral infections, fixed in our genome and transmitted vertically [5]. Integration of HERVs has occurred over millions of years and over time they have acquired point mutations rendering them unable to produce infectious virions [6,7]. The youngest retrovirus in the HERV family is the betaretrovirus HERV-K (HML-2), which was the most recent HERV to integrate into the genome and thus has the highest degree of transcriptional activity [8]. Furthermore, of the 31 defined HERV families [9], the sequence of HERV-K is the most homologous to HIV and is the most widely examined HERV in the field of HIV [10]. The relationship between HERV-K and HIV emerged from initial reports demonstrating that antibodies against HERV-K (HML-2) were found in 70% of HIV-positive patients compared with only 5% of HIV-negative individuals [6,11]. Furthermore, it has been shown that HERV-K RNA titers are elevated in the plasma of HIV-infected individuals in comparison with seronegative individuals [3,12,13,14]. Most intriguing for vaccine design is a recent study demonstrating elevated and functional T-cell responses in HIV-positive patients against cellular protein targets derived from HERV-K [3]. Given that HERVs are encoded in the germ-line, it is hypothesized that they are not subjected to the same degree of cytotoxic T lymphocyte (CTL) immune pressure and subsequent immune escape as HIV antigens. In addition, plasma HERV-K RNA titers are inversely correlated with the degree of HIV suppression following highly active antiretroviral therapy (HAART) suggesting that HERV-K may be a predictor of HIV replication [12], although the mechanism by which HIV impacts HERV-K transcriptional processes is not well understood.

Many studies have examined HERV-K expression in the plasma of HIV-infected patients; however, little work has been
done to determine if HERV-K is activated at the cellular level. It is known that peripheral blood mononuclear cells (PBMCs) express HERV-K gag and pol RNA [15]. The limitations of many human in vivo studies are that these findings are derived from independent cross-sectional samples. A longitudinal study using a non-human primate model will allow for endogenous retrovirus (ERV) expression to be assessed pre- and post-SIV infection from dependent samples, providing a precise measurement of changes in ERV expression as a result of viral infection.

Cynomolgus macaques (Macaca fascicularis) are a widely utilized non-human primate model for preclinical biomedical research [16]. This species of macaques is susceptible to pathogenic simian immunodeficiency virus (SIV) infection and ultimately succumbs to simian AIDS (SAIDS). Although the duration of SIV pathogenesis and progression to SAIDS is shorter in macaques than in humans [17], cynomolgus macaques closely recapitulate HIV disease progression and thus are a practical model to assess endogenous retrovirus expression. In this study, we evaluated cynomolgus macaque endogenous retrovirus (CyERV) expression in PBMCs following SIV infection. Here we show the characterization of CyERV genes and the effect of vaccination and SIV infection on CyERV expression levels in PBMCs.

Results
Isolation and Genetic Characterization of CyERV Genes
Full-length CyERV envelope and gag genes were PCR-amplified and sequenced to examine the diversity between CyERVs as well as to determine if cynomolgus macaque endogenous retroviruses encode intact open reading frames (ORFs). With respect to nucleotide identity, the isolated CyERV clones showed a high degree intra-animal homology for env (avg. 90% identity) and gag (avg. 94% identity), and inter-animal homology for env (91% identity; n = 5) and gag (81% identity; n = 4). A BLAST search of the CyERV env and gag sequences revealed homology to HERV-K, while homology to other known HERV families was not observed. The amino acid length (avg. 695 aa) and molecular weight (avg. 75 kDa) of the CyERV genes were comparable to HERV-K genes (env 694 aa, 78 kDa and gag 666 aa, 74 kDa) [18]. The majority of the CyERV env and gag amino acid sequences contained multiple stop codons likely as a result of acquired point mutations. In this regard, the CyERV env clones were of two distinct genotypes. Of the 10 CyERV env clones (2 clones/animal) isolated, half represented genotype 1 (19 stop codons; 696 aa) and half represented genotype 2 (29 stop codons; 691 aa), while each genotype shared 99.4% nucleotide identity between clones. From our analysis, these stop codons are likely the result of point mutations with genotype 1 resulting from 7 indels and 4 substitutions while genotype 2 was derived from 9 indels and 5 substitutions. To evaluate the amino acid identity between CyERV env and HERV-K env (GenBank: CAA68686), we resolved the indels/substitutions in the CyERV sequence and determined that genotypes 1 and 2 share 86.5% and 87.9% amino acid identity, respectively, with HERV-K env. The CyERV gag isolates contained an average of 27 stop codons (range: 0–47), however one isolate encoded a complete CyERV gag ORF (GenBank: JN993533) with 695 amino acids and a predicted molecular weight of 77.7 kDa. This complete CyERV gag clone shares 77.4% identity with HERV-K102 gag (GenBank: P63130) at the amino acid level (Fig. 1). The discovery of an intact CyERV gag ORF supports previous documentation that cynomolgus macaques may have retained the capacity to produce full-length CyERV gag proteins [19].

Impact of Vaccination on CyERV Transcript Expression Levels in PBMCs
The SIV vaccine trial was comprised of two vaccine groups (Δ5-CMV and Δ6-CCI) and one control group (mock-immunized), and followed a prime-boost-boost immunization regimen. The animals in the vaccine groups were immunized with highly attenuated SIVmac239 viral constructs, as described in the Materials/Methods and Figure 2 [20]. To assess the impact of vaccination on CyERV gene expression levels, we extracted RNA from PBMCs isolated at baseline and post-vaccination timepoints to quantify CyERV env and gag gene expression by RT-qPCR. The post-vaccination timepoint is derived from samples taken at either week 114 or 116. No significant differences were observed between the two sampling weeks and as such these post-vaccination samples were grouped. There were no differences in CyERV env and gag expression between the three vaccine groups (controls, Δ5-CMV, and Δ6-CCI). In comparison with baseline expression (log10 −3.30), CyERV env transcript levels were significantly lower following vaccination (log10 −4.03; n = 6; P = 0.01) (Fig. 3A), with a less pronounced decrease in CyERV gag expression between baseline (log10 −2.53) and post-vaccination (log10 −3.04; n = 6; P = 0.33) (Fig. 3B).

Downregulation of CyERV Transcript Levels Following SIV Infection in PBMCs
Following vaccination, the animals were challenged weekly with a multi-low-dose SIVmac239 intrarectal challenge regimen until productive plasma infection was detected by RT-qPCR [20]. We examined CyERV env and gag gene expression at two timepoints representing acute (avg. 6 weeks post-infection (wpi); range: 4–10 wpi) and chronic (avg. 79 wpi; range: 25–98 wpi) SIV infection (Fig. 2). CyERV gene expression was significantly decreased during acute SIV infection and returned to baseline levels during chronic SIV infection (Fig. 3). At the acute timepoint, the mean CyERV env expression level (log10 −4.74) was lower than baseline (log10 −3.24; n = 10; P < 0.001) (Fig. 3A). CyERV env transcript expression increased during chronic SIV infection compared with post-vaccination (log10 −3.65 versus log10 −3.96; n = 10; P = 0.03) and acute SIV infection (log10 −3.61 versus log10 −4.55; n = 11; P = 0.01). Mean CyERV gag expression decreased following acute SIV infection (log10 −3.40) compared with baseline levels (log10 −2.40; n = 7; P = 0.02) and was comparable with baseline during chronic SIV infection (Fig. 3B). Of note, when the above data was normalized with respect to CD4+ T-cell counts, CyERV gene expression patterns were analogous (data not shown).

Susceptibility to SIV Infection was Associated with CyERV Expression Levels in PBMCs
Susceptibility to SIV infection was determined based on the number of SIVmac239 challenges required to infect the cynomolgus macaques during the challenge phase of the trial [20]. Consistent with the grouping assigned by Willer et al. [20], the groups were based on the number of SIV challenges (<15 vs. ≥15 challenges) to infection, irrespective of vaccination or control group. CyERV env gene expression levels in the animals that required ≥15 SIV challenges (n = 5) to establish infection were elevated across all timepoints compared with the animals that were infected with SIV in <15 challenges (n = 7) (Fig. 4A), although there was no statistically significant difference between the two groups. Likewise, CyERV gag expression in the animals that required ≥15 SIV challenges to establish infection was increased across all timepoints except during chronic SIV infection where gag expression levels were slightly lower compared with animals...
that were infected with SIV in <15 challenges (Fig. 4B). CyERV gag transcript levels were significantly higher at the post-vaccination timepoint in the animals that were less susceptible to SIV infection ($\log_{10} 2^{2.23}$) compared with those that were more susceptible ($\log_{10} 2^{3.47}$; $P = 0.02$).

CyERV Gene Expression Associates with CD4$^+$ T-cell Counts and SIV Viral Load

Given that acute SIV infection had the most significant impact on CyERV expression levels, we examined the relationship between CyERV expression and clinical markers of disease status using CD4$^+$ T-cell counts (CD4$^+$ T-cells/μL of PBMCs) and SIV viral load ($\log_{10}$ viral load/mL), previously determined during the vaccination study [20]. CyERV gene expression negatively correlated with CD4$^+$ T-cell counts during acute SIV infection for env ($r = -0.54$; $n = 11$; $P = 0.08$) and gag ($r = -0.74$; $n = 10$; $P = 0.02$) (Fig. 5A-B). Additional correlations were performed using the percent CD4$^+$ T-cells of total lymphocytes and the results were comparable to the absolute counts (data not shown). With respect to SIV viral load, CyERV env and gag gene expression did not correlate significantly with viral load.
not significantly correlate with acute SIV viral load. CyERV Env expression ($r = -0.17; n = 9; P = 0.67$) showed a trend towards a negative association (Fig. 5C) in comparison with gag expression ($r = 0.11; n = 8; P = 0.80$), which displayed a trend towards a positive association (Fig. 5D). A similar analysis of CyERV expression was conducted during chronic SIV infection with no significant correlations observed with respect to either CD4$^+$ T-cell counts or SIV viral load (data not shown).

CyERV Transcript Expression in in vitro SIV-infected CD4$^+$ T-cells

To examine CyERV gene expression in a single population of cells, CD4$^+$ T-cells were enriched from cynomolgus macaque PBMCs and mock- or SIV-infected (SIVmac239) by magnetofection [21]. Total RNA was isolated from the mock- or SIV-infected cells to quantify CyERV env and gag expression by RT-qPCR. Although, hypothetically, all cell types have the potential to produce CyERVs, we specifically examined CD4$^+$ T-cells since they are the cell type that is preferentially infected by SIV. CyERV transcript levels were examined in mock- and SIV-infected CD4$^+$ T-cells to evaluate the effect of SIV infection on CyERV gene expression. CyERV env expression was not significantly altered in SIV-infected CD4$^+$ T-cells compared with the mock-infected CD4$^+$ T-cells ($n = 5; P = 0.16$) (Fig. 6A). Eighty percent (4/5) of the animals showed an increase in CyERV gag expression following SIV infection; however, no statistically significant difference was observed between mock- and SIV-infected CD4$^+$ T-cells ($n = 5; P = 0.31$) (Fig. 6B). No correlation was observed between CyERV expression and the number of SIV-infected CD4$^+$ T-cells (data not shown).

Discussion

It is unclear how many CyERV proviruses have been incorporated into the cynomolgus macaque genome and whether they remain intact. The rhesus macaque (*Macaca mulatta*) genome contains nineteen complete proviruses of rhesus macaque endo-
genous retrovirus-K (RhERV-K) [22]; however, a similar analysis has yet to be performed in the two recently sequenced cynomolgus macaque genomes [23,24]. Given that these Old World monkeys (OWMs) evolved from a common ancestor after the integration of HERV-K [25,26], we would expect cynomolgus macaques to harbor a similar number of CyERV-K proviruses in their genome.

Figure 5. CyERV gene expression negatively associated with CD4$^+$ T-cell counts and SIV viral load. During acute SIV infection, CyERV envelope (n = 11) and gag (n = 10) gene expression levels negatively correlated with CD4$^+$ T-cell counts (CD4$^+$ T-cells/μl of PBMCs) (A-B) and negatively associated with acute SIV viral load for CyERV envelope (n = 9) (C). CyERV gag (n = 8) expression showed a slight trend towards a positive association with acute SIV viral load (D). Bivariate correlations were performed using Pearson’s correlation coefficient (r), statistical significance values (p) are shown. doi:10.1371/journal.pone.0040158.g005

Figure 6. CyERV gene expression levels in CD4$^+$ T-cells following in vitro SIV infection. CyERV gene expression was quantitated by RT-qPCR using RNA isolated from mock- and SIV-infected CD4$^+$ T-cells. CyERV envelope (A) and gag (B) gene expression levels relative to GAPDH are compared between in vitro mock- and SIV-infected CD4$^+$ T-cells. Paired-samples t-tests (2-tailed) were performed to determine statistical significance. doi:10.1371/journal.pone.0040158.g006
Furthermore, after the divergence of hominoids and OWMs 25 million years ago, there have been at least 2730 OWM-specific ERV sequences identified and it is possible that these insertions may be more active [26]. A study examining the CyERV families and the number of proviruses present in the cynomolgus macaque genome would aid in elucidating the full CyERV complement and direct future CyERV-based SIV/HIV vaccine approaches. It is currently unclear if SIV selectively alters CyERV gene expression in certain families or if the phenomenon is universal. Our documentation of a full-length CyERV gag ORF (Fig. 1) indicates that cynomolgus macaques have retained the capacity to produce CyERV gag proteins, as previously described in a study examining gag sequences in OWMs [19]. Although we sequenced a number of CyERV env clones we were unable to isolate an intact ORF, supporting previous observations by Mayer et al. [19]; however, we cannot rule out the potential for full-length ORFs to be present. Our analysis was directed at characterizing CyERV env and gag, therefore it is possible that other intact CyERV ORFs (such as pol) are present in the cynomolgus macaque genome.

In the current study, CyERV expression levels were evaluated in PBMCs from a cohort of cynomolgus macaques utilized in a live attenuated SIV vaccine trial. As a protracted immunization was employed [20], we examined its impact on CyERV gene expression. Our results indicated that following the vaccination protocol there was a significant decrease in CyERV env expression levels and a marginal decrease in CyERV gag expression (Fig. 3). This was consistent between vaccinees and controls suggesting that part of the alterations in CyERV gene expression. In a recent study examining CyERV expression in the brains of cynomolgus macaques infected with bovine spongiform encephalopathy (BSE), the authors suggested that the observed decrease in macaque ERV-K (HML-2) gag RNA and protein expression may have been attributed, at least partially, to the inoculation procedure [27]. Ideally, to assess the impact of SIV infection on CyERV gene expression, a cohort of naive animals that have not previously been vaccinated would be the preferred model. Of note, a comparison of CyERV gene expression levels with respect to the use of CpG as an adjuvant showed no effect on CyERV gene expression levels (data not shown). As such, it is possible that innate immune responses, although not mediated through CpG, may be playing a role in downregulating CyERV gene expression.

During acute SIV infection, CyERV env and gag gene expression levels were the lowest in comparison with all other timepoints (Fig. 3). If innate immunity is involved, maximal increase in innate immune activation during acute SIV infection could be contributing to the decrease in CyERV transcript levels. Preliminary data from our group examining the effects of Varicella Zoster Virus (VZV) infection on CyERV transcript expression showed similar results with decreased CyERV env and gag expression levels during acute VZV infection and a return to baseline levels during chronic VZV infection (unpublished data), further suggesting that the strong innate immune response during acute infection may be involved in decreasing CyERV gene expression.

Our observations that demonstrated disparate expression levels with respect to SIV susceptibility may be explained by a concept examined by Garrison et al. [5] in which they suggest that HERV-K-specific T-cells may cross-react with HIV-specific epitopes. In a separate study, they showed HERV-K-specific T-cells might be involved in the control of HIV during chronic infection [4]. In our study, animals that required a greater number of challenges with SIV to establish infection had higher CyERV env and gag transcript levels prior to SIV challenge and during acute SIV infection (Fig. 4). Thus, the positive association between CyERV expression levels and reduced susceptibility to SIV infection might imply that CyERV-specific cytotoxic CD8+ T-cells recognize SIV-specific epitopes thereby killing SIV-infected cells and contributing to the prevention of SIV infection. However, once SIV infection occurs our results show a trend towards a negative association between CyERV env expression levels and SIV viral load (Fig. 5C).

This challenges the hypothesis that CyERV gene expression increases in SIV-infected cells. Furthermore, a negative correlation was observed between CyERV env and gag expression and CD4+ T-cell counts during acute SIV infection (Fig. 5A-B); whereas, during chronic SIV infection there was a trend towards a positive association between CyERV gene expression and CD4+ T-cell counts. Considering that CD4+ T-cells are a major target for SIV infection, we would expect to see the most robust effect of SIV infection on CyERV transcript levels in this subset of cells.

We examined in vitro SIV-infected cynomolgus macaque CD4+ T-cells to evaluate whether CyERV expressions was modulated by SIV infection. SIV infection had no impact on CyERV env expression; however, four out of five animals showed increased CyERV gag expression in the SIV-infected CD4+ T-cells (Fig. 6). Recently, Lefebvre et al. [28] used next-generation sequencing to examine cellular transcription of HERVs in an HIV-infected T-cell line and showed an significant increase in HERV-K transcription in the HIV-infected cells. Although the methodology differed and their analysis was not specific to any one particular HERV protein, these similar results would suggest that intracellular endogenous retrovirus transcriptional events do not parallel the degree of upregulation observed by others examining HERV-K expression in HIV-infected plasma samples. Although plasma samples from this study were not available, matched sets of plasma and PBMCs would be useful in delineating any differences in HERV expression between the two sources.

Our findings from the evaluation of CyERV expression in PBMCs isolated from SIV-infected cynomolgus macaques were not consistent with what has been observed for HERV-K expression in the plasma of HIV-infected humans. During acute SIV infection, CyERV transcript levels were decreased and negatively associated with CD4+ T-cell counts. Future examination of CyERV expression in plasma and tissue samples would assist in further explaining whether these findings are unique to macaques or if this is a cellular phenomenon that has yet to be fully investigated in humans. Based on the results from this non-human primate study, an investigation into the direct effects innate immune activation and SIV infection on CyERV expression will be required before CyERVs can be evaluated as SIV vaccine targets.

**Materials and Methods**

**Ethics Statement**

The samples in this study are historical samples derived from a previous SIV vaccine trial [20] and no further samples were collected for the current study. All animal work for the original SIV vaccine trial was for research purposes and was approved in accordance with the Health Canada Institutional Animal Care Committee (protocol #2010-001), which met the ethical, scientific and social responsibility criteria set out by the Canadian Council on Animal Care ([http://www.ccac.ca/Documents/Standards/Guidelines/Protocol_Review.pdf](http://www.ccac.ca/Documents/Standards/Guidelines/Protocol_Review.pdf)). For the SIV vaccine trial study, adult male colony-bred cynomolgus macaques (Macaca fascicularis) of Philippine origin were housed at the Animal Resources Division at the St. Frederick Banting Research Center (Ottawa, Canada).
The veterinary staff monitored the animals daily for food intake, stool consistency and general welfare. The animals were anesthetized with ketamine (10 mg/kg) during all inoculations, viral challenges and sampling procedures. Any abnormal observation found during regular clinical evaluations was brought to the attention of veterinary staff. In situations of multiple systemic consequences, coupled with complete anorexia for more than three days resulting in weight loss of more than twenty percent, euthanasia was elected while the animal was under palliative care/feeding. The decision to euthanize the animal was always to ensure the animal would not suffer, and at no time was death considered an acceptable endpoint. Animals had large single cages excluding the minimum requirements with areas of both privacy and visual social interaction with other animals. Animals were given a daily comprehensive program of environmental enrichment to prevent abnormal behavior and minimize stress. On alternate days, animals were given access to a large exercise area.

**Animals and Vaccine Constructs**

All samples in this study are derived from a previous SIV vaccine trial [20], in which twelve adult male cynomolgus macaques of Philippine origin were randomly assigned into two experimental groups and one control group (4 animals in each group). Two highly attenuated vaccine constructs (Δ5-CMV and Δ6-CCI), derived from SIVmac239 [29,30], were employed in the SIV vaccine trials. The vaccination schedule followed a prime-boost regimen in which the animals were primed with viral constructs (Δ5-CMV or Δ6-CCI) or medium only (control animals), and subsequently boosted with plasmid DNA (Δ5-CMV, Δ6-CCI or control). For the second boost, the animals were given their respective viral and plasmid constructs (Δ5-CMV or Δ6-CCI), while the controls received medium alone and control plasmid DNA. All animals received a B-class CpG oligodeoxynucleotides (ODN) adjuvant with their viral construct and/or plasmid DNA during the second boost [20]. SIV inoculations were repeated weekly until a detectable SIV infection was established. Peripherally blood mononuclear cells (PBMCs) were isolated using routine methods [20] at various stages throughout the trial and stored at −150°C. Samples from the following timepoints were used for the current study: baseline (pre-vaccination), post-tutions were as follows: 95

**Evaluation of CyERV Expression Post-SIV Infection**

CD4+ T-cell Enrichment and in vitro SIV Infection

CD4+ T-cells were enriched from cynomolgus macaque (n = 5) PBMCs using a custom EasySep negative selection kit (StemCell Technologies). Enriched CD4+ T-cells were resuspended in R15–100 buffer (RPMI 1640 plus 15% FBS, 1-glutamine, antibiotic/antimycotic) supplemented with 100 IU/ml rIL-2 and 2 μg/ml concanavalin A and cultured in vitro for 3–6 days. The CD4+ T-cells were either mock- or SIV-infected (SIVmac239) by magnetofection [21] for 48 hrs and stained with SIVmac p27 antibody (NID AIDS Research and Reference Reagents Program) to evaluate SIV infection by FACS. Total RNA was isolated from the cell pellets.

**RNA Isolation**

PBMCs were thawed at 37°C, washed with DMEM (Sigma-Aldrich) supplemented with 10% FBS (Wisent Bioproducts), 100 U/ml penicillin, 100 μg/ml streptomycin (Sigma-Aldrich), and subsequently washed with 1X Dulbecco’s phosphate buffered saline (Invitrogen). Total RNA was purified from the cell pellets using RNAasy Plus Mini kit (Qiagen) and eluted in 30 μl of RNase-free water, as per manufacturer’s protocol. The RNA was treated with 2 U TURBO DNase (Ambion) to remove any contaminating genomic DNA.

**Standard Curve Plasmid Cloning**

A control plasmid for quantitative polymerase chain reaction (qPCR) was generated by incorporating the ampiclons of CyERV env and gag, in conjunction with a reference gene, cynomolgus macaque glyceraldehyde 3-phosphate dehydrogenase (GAPDH). Restriction sites were incorporated by PCR to facilitate cloning of the genes into pMECA cloning vector (GenBank: AF017063). CyERV env primers were designed in regions conserved between the two clones of CyERV env, to amplify a 166 bp amplicon, FP 5′-AGGTGCGGCGGCCTACCTGCGCCCAAGTGAC-3′ (Nol underlined) and RP 5′-ACGTGGCTGTGGGCTCGGTCAACCAATTTGTG-3′ (Nol underlined). CyERV gag primers were designed in conserved regions between the two clones of CyERV gag, to amplify a 301 bp amplicon, FP 5′-AGCTGGAATCAGCTGAGTACAGTTCATCGAG-3′ (Nol underlined) and RP 5′-AGCTGATACCGTAGTCAATGCCTGCAATTCATG-3′ (Bgl II underlined). The GAPDH primers were designed from a known *Macaca fascicularis* mRNA partial coding sequence (GenBank: DQ466111). PCR primers ( FP 5′-ACGGCCATGTGACCTGGCGCTGTCGAAAG-3′ [SpI underlined] and RP 5′-ACGTGCTTGAGGTCCTCGAGGCCTGCTTTCA-3′ [NheI underlined]) were designed to amplify an 80 bp region of GAPDH from cynomolgus macaque PBMCs. All gene fragments were cloned into pMECA (CyERV env; CyERV gag; GAPDH) and confirmed by Sanger sequencing. Serial dilutions of the plasmid were prepared to achieve a linear range of 8 logs (5 to 5^9) and single use aliquots were stored at −80°C.

**Reverse Transcription-quantitative Polymerase Chain Reaction (RT-qPCR)**

CyERV mRNA transcript levels were assessed by RT-qPCR. Total RNA (avg. 324 ng) was reverse transcribed to complementary DNA (cDNA) using SuperScript III First-Strand Synthesis SuperMix (Invitrogen) as per manufacturer’s protocol. Reverse-transcription negative controls and no template controls were included for each sample. All samples were loaded as technical triplicates. Thermal cycling conditions were: 95°C for 10 mins, 40 cycles of 95°C for 15 s, 60°C for 1 min followed by a dissociation
stage of 95°C for 15 s, 60°C for 15 s and 95°C for 15 s. The primer sequences and amplicon sizes of each gene are listed in Table 1. Representative qPCR amplicons were analyzed by gel electrophoresis on a 2% agarose gel to confirm the size and purity of the products. In addition, selected qPCR products were sequenced to confirm the transcripts. The RT-qPCR assay consistently demonstrated high efficiency and linearity for each gene (Table 1). A standard curve was used to enumerate CyERV mRNA expression. Sequence Detection System v2.4 (Applied Biosystems) was used for gene expression analysis with CyERV expression levels being standardized to 1 μg of RNA, normalized to the GAPDH reference gene and log transformed. Values represent the mean of the three technical replicates. Threshold cycle values or quantitative values that were deemed outliers based on Grubbs’ test (Z ≥ 1.5) were excluded.

Statistical Analysis
One-way ANOVA tests were performed to compare the mean differences between the vaccine groups (Δ5-CMV, Δ6-CCI, control). Paired-samples t-tests (2-tailed) were performed to determine the statistical significance among various timepoints (dependent samples). For inter-group comparisons, independent-samples t-tests (2-tailed) were applied to determine the statistical significance. Bivariate correlations were performed using Pearson’s correlation coefficient (2-tailed). Statistical analysis was completed using SPSS for Macintosh (Rel. 19.0.0. 2010. SPSS Inc.).

Acknowledgments
We would like to thank Dr. R. Brad Jones & Dr. Mario Ostrowski (University of Toronto, Toronto, Ontario, Canada) for their technical expertise and review of the manuscript. We would like to thank Dr. Paul Sandstrom & Richard Flom (National HIV and Retrovirology Laboratories, National Microbiology Laboratory, Public Health Agency of Canada, Ottawa, Ontario, Canada), Dr. Joseply Fournier (Scientific Services Division, Health Products and Food Branch, Health Canada, Ottawa, Ontario, Canada) and Dr. Mark A. Wainberg (McGill University AIDS Centre, Montreal, Quebec, Canada). We sincerely thank the veterinary and technical staff at the NHP colony of Health Canada. We thank Dr. Nico Nagelkerke for his assistance with the statistical analysis.

Author Contributions
Conceived and designed the experiments: AKM DOW KSM. Performed the experiments: AKM OS OHI JKC. Analyzed the data: AKM DOW KSM. Contributed reagents/materials/analysis tools: KSM. Wrote the paper: AKM DOW KSM.

Table 1. RT-qPCR Primers.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer Sequence</th>
<th>Amplicon</th>
<th>Slope</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH</td>
<td>F: TGACCTGGCCTGCTGGAAAA</td>
<td>68bp</td>
<td>−3.38</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>R: CTCGGAGCCCTGTCTCA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CyERV Envelope</td>
<td>F: TACCTGGGCCCCACAGATGAC</td>
<td>88bp</td>
<td>−3.41</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>R: GGAGGATATGGATACCCAGTGGA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CyERV Gag</td>
<td>F: GGCCCTGGGAAAGAATCCCAAG</td>
<td>289bp</td>
<td>−3.55</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>R: TGCATATCCCTGCCATCCATC</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The slope and correlation represent the mean value. doi:10.1371/journal.pone.0040158.t001

The slope and correlation represent the mean value. doi:10.1371/journal.pone.0040158.t001

References

References


RNA titers in HIV-1-infected patients receiving successful versus unsuccessful highly active antiretroviral therapy. AIDS Res Hum Retroviruses 23:1083-1086.


377. **IAVI.** 2013. Clinical Trials Database.


222


HIV viral set point and host immune control in individuals with HIV-specific CD8+ T-cell responses prior to HIV acquisition. AIDS 24:1449-1454.


the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. J Virol 68:4650-4655.


Ohtaki, S., R. Hondo, H. Kodama, and T. Kurata. 1988. Experimental activation of latent cytomegalovirus infection of the captive bred (F1) cynomolgus monkeys by live or


242


701. **Rowe, P. M.** 1996. A cofactor for HIV-1 entry into cells is identified. Lancet **347:**1395.


837. UNAIDS. 2011. Implementation of the Declaration of Commitment on HIV/AIDS and the Political Declaration on HIV/AIDS.


of human immunodeficiency virus-type 1 during the natural course of infection. Science 272:537-542.


